UK Migration at the Beginning of the 21st Century

Vincent Judson

Master of Philosophy

Heriot Watt University

School of Social Sciences

June 2017

Abstract

This work analyses different aspect of migration in between the first two censuses that were conducted in the UK in the twenty-first century. The second chapter examines the factors that can have an effect on the rate at which migrants in England and Wales leave the UK in order to return to their home country during the time period between the census of 2001 and that of 2011. It uses a theory, initially developed by Borjas and Bratsberg (1996), that attempts to explain why individuals from countries outside of the UK may choose to leave the country after their arrival. Leaving may have been pre-planned before the initial migration to the host nation or may be the result of poor outcomes after arrival. It is found that many of the results found by Borjas and Bratsberg (1996) in the United States are replicated in the first decade of the Twenty-first century for the United Kingdom. In the third chapter I investigate the reasons into why there is a wage gap between native and migrant workers in the UK labour market. It provides evidence to show that the gap in wages between the two populations can be linked to the composition of the migrant cohort. It also shows that as the European Union expanded there was a significant change in the discrimination against the migrant from the countries that joined the EU in May 2004 and January 2007.

Dedication

I would like to thank everyone who has helped me to recover after my diagnosis with Multiple Sclerosis in 2012. This includes, but is not limited to friends, colleagues, staff at the University, my Supervisors on this work and all of my family.

Table of Contents

1. Introduction 1

2. Return Migration from the Host Nation 8
2.1 Introduction 9
2.2 Literature Review
2.2.1 Who Returns? The Outmigration of the Foreign Born (1996) 10
2.2.2 Other Relevant Papers 12
2.3 Model 18
2.4 Data 22
2.5 Methodology 24
2.6 Results 27
2.7 Conclusion 31

3 Are Migrants Discriminated Against in the UK Labour Market? An Oaxaca-Blinder Wage Decomposition. 33
3.1 Introduction 34
3.2 Literature Review 34
3.2.1 A Short Review of the Literature on Discrimination 34
3.2.2 A Selective Review of the Wage Decomposition Literature 36
3.3 Model 44
3.4 Data 44
3.5 Methodology 47
3.6 Results 51
3.6.1 All Migrants and all Natives
3.6.1.1 OLS Results 51
3.6.1.2 Twofold Decomposition 51
3.6.1.3 Threefold Decomposition 55
3.6.1.4 Counterfactuals 57

3.6.2 Male
3.6.2.1 OLS Results 57
3.6.2.2 Twofold Decomposition 58
3.6.2.3 Threefold Decomposition 59
3.6.2.4 Counterfactuals 59
3.6.3 Female
3.6.3.1 OLS Results 60
3.6.3.2 Twofold Decomposition 60
3.6.3.3 Threefold Decomposition 61
3.6.3.4 Counterfactuals 62
3.7 Conclusion 63
5 Conclusion 65
4 Appendix A 68
5 Appendix B 74
Bibliography 99

Tables

Chapter Two Tables

Table 2.1 Descriptive Statistics 23

Table 2.2 Regression upon In and Out migration rate 69

Table 2.3 Marginal Effects of Independent Variables 70

Chapter Three Tables

Table 3.1 Descriptive Statistics for Natives and Migrants

Table 3.1.1 Total Population 75

Table 3.1.2 Male Natives and Migrants 76

Table 3.1.3 Female Natives and Migrants 77

Table 3.2 Occupational Distribution of the Different Groups 78

Table 3.3 Percentage that have Migrated from Various World Regions 79

Table 3.4 OLS Results for Natives and Migrants (Total Populations) 80

Table 3.5 Twofold Decomposition with Natives as the Non-discriminatory Group 81

Table 3.6 Threefold Wage Decomposition 82

Table 3.7 OLS Results for Male Sample of Natives and Migrants 83

Table 3.8 Twofold Decomposition for Male Natives and Migrants with Natives as Non-discriminatory Group 84

Table3.9 Threefold Wage Decomposition for Male Natives and Migrants 85

Table 3.10 OLS Results for Female Natives and Migrants 86

Table 3.11 Twofold Decomposition for Female Natives as the Non-discriminatory Group 87

Table 3.12 Threefold Wage Decomposition 88

Table 3.13 Predicted Average Wage Counterfactuals 89

Figures

Chapter Three

Figure 3.1 Wages of both Migrants and Natives 49

Appendix A

Figure 2.1 Positive Selection 71

Figure 2.2 Negative Selection 72

Figure 2.3 World Population Inflow to the UK 73

Chapter One

Introduction

In this thesis I study the outcomes that are achieved by the migrant population in the United Kingdom of Great Britain and Northern Ireland (UK) in the period between the censuses in 2001 and the next census that was conducted in 2011. In this work, I primarily answer two questions about the migrant population within the UK. The first question asks what proportion of the migrant population, who arrived during this period, remain at the time of the census in 2011 was conducted. I next ask what the difference in earnings is between the two populations are at various points in time during the study period. The results that are found in the second chapter would indicate that almost two-thirds of those who came to the UK over the ten year period have left by the time that the census information was gathered in 2011. I discover, in the third chapter, that there is always a gap in the earnings between the two groups. In 2001 it was observed that the migrant population was earning more than the native population. However, by 2011 the gap has changed from migrants earning more in 2001, on average, to the native cohort being the group that is earning more by 2011.

Two fundamental definitions need to be specified here as they are used throughout this work. The first of these is the host country. A host country is a country that migrants arrive to reside in. This may be on a temporary basis or for permanent resettlement. The main host country in this work is the UK.[1] Source countries could also be termed as countries of origin. These are the countries where a migrant originates from and where they currently hold nationality. Figure 2.3, on page 79, shows a map of where the source nations included in this analysis are located.

As mentioned above the second chapter tends to agree with the original paper by Borjas and Bratsberg (1996). They find that many of the migrants actually move on after a number of years. This may be back to their source country or onto a third country. In this work, like in the work by the original authors, I assume that if a migrant is no longer counted as being a resident in the UK then they have migrated back to their source country.

I explore various reasons that lie behind the decision that the migrants have made to return to their source country. The work in Chapter Two finds that the main reasons for this decision being made are the cost of migration and the disparity in the level of wealth between the host and source nations.

[1] Other host countries are mentioned but these are in works by other authors.

The third Chapter then investigates if there is a difference in the earnings between two well-defined groups. These two groups are one group of the native population and another group of migrants who are now resident in the UK. I examine if a gap between the mean earnings of the two groups exists. If such a gap is present, the econometrics that I am using will be able to point to where the gap is being derived from. I use an Oaxaca-Blinder wage decomposition in order to investigate if the migrant cohort is paid less than the native cohort is remunerated by. If one group is paid less than the other is, then the wage decomposition will be able to show the reasons why such a gap in earnings may exist.

Furthermore, Chapter Two aims to examine the factors that have an effect on the rate at which migrants in England and Wales leave the UK in order to return to their source country between the two UK censuses. It uses the theory that was developed by Borjas and Bratsberg (1996). There are a number of reasons for leaving the source nation. It could be that returning was pre-planned before the migrant had ever left the source country for their initial migration, or it could be the result of changing circumstances in the host or source nation, or even in both. If the gap between the source and host nations has decreased then it would be expected for more migrants to want to return to their country of origin. It could also be that after arrival, perhaps due to poor information or bad luck that a migrant may not be able to achieve what they expected to and therefore have returned.

This study then estimates how much change, in a variety of factors specific to the economic position of the source country in relation to the host nation that the migrant is residing in at the time the survey was conducted, may lead to a migrant leaving the host nation. I did this by applying a grouped probit model to the data that I have collected. I also apply marginal effects at various means to get an estimated number of individuals who will engage in this type of behaviour.

I employ a number of variables to run this regression and see which of the variables has the largest effect on the re-migration rate. These variables include the gap in GDP between the source (or home) country and the host nation. I also included, as explanatory variables, the distance between the capital of the source country and London, the income inequality that is present in the source country along with the square of this variable and the political stability in the source country. The empirical

results of this regression indicates that there has been a change in the distribution of migrants source countries between the two censuses. However, these results do not tell the reader about the size of any change in the selection that is produced by the migration flows.

I examine the numbers that produce these changes in the size of migration flows and use marginal effects at the mean in order to investigate these. The marginal effects of the various migrant populations are used in order to estimate the number of individuals who are engaging in migration, both arriving in and leaving from the host nation. I produce three different types of marginal effects. The different types of marginal effects that are employed in this chapter are the marginal effect at the mean for all countries, one done solely for India, the largest migrant group in 2001, and one done for Poland, which was the largest inflow group after 2004.

This leads to the question of why so many individuals have chosen to migrate to the UK, and also then why they choose to return to their country of origin. A variety of factors will attract migrants to a particular host nation. Changes in these variables may lead to the opposite effect, those initial migrants deciding to return to their source country. This chapter then seeks to identify those variables that draw the migrants to reside in the UK and then to return to their country of origin. A more important aspect of migration is the type of individual that decides to migrate to and to make the UK their place of residence. If the average migrant from a particular source country is positively selected then they would derive from the upper half of the distribution of skills in the country that they came from. The opposite is true if these are negatively selected migrants.

In the chapter, the work of Borjas and Bratsberg (1996) is applied to the UK labour market. These authors find that many migrants were not migrating to the US to become permanent residents but were moving there for a short period before migrating back to their source country. Here I find that there is a similarity to the pattern that was found by Borjas and Bratsberg (1996). However the overall numbers of individuals is much larger due to the reduced costs of travel.

I observe that the selection that results from the return migration is dependent upon the selection of the initial cohort from a source country that undertook migration to the host nation. To proxy for the cost of migration I use distance, and that is found to be one of

the most important factors affecting the migration choices that are made by individuals. The marginal effects for distance for the overall group show that if there is an extra thousand miles was added to the distance that had to be travelled then 699 more individuals are going to leave the UK in order to return to their source country. A similar effect is observed if there is an increase in GDP with 1102.64 more people returning to their source country.

The return migration just exacerbates the selection that defined the original cohort. This means that those migrants who remain can be termed as the "best of the best" or the "worst of the worst" in those remaining individuals are from the tails of the distribution in skills. Anyone who is in between those two extremes is more likely to go back to their country of origin.

Chapter Three then investigates the reasons why there is a wage gap between the average worker in both the native and migrant groups in the UK labour market. In this chapter, all of the migrant groups from the previous chapter are amalgamated into a group that is defined as being all of the migrants to the UK and are studied against a representative group of native individuals. It provides evidence to show the gap in wages between the two populations can be linked to the composition of the migrant cohort. It also shows that as the European Union expanded there was a significant change in the wage structure that was caused by individuals from the accession states that joined the EU in May 2004 and January 2007 arriving in the UK labour market.

The analysis that will be performed in this chapter will be done by using the work Oaxaca-Blinder wage decomposition technique that was developed by Ronald Oaxaca (1973) and Alan Blinder (1973) and shows discrimination through distinct differences in both the returns to education and the returns to tenure that the migrant group receives, relative to the native cohort. The data comes from the UK Labour Force Survey (LFS) for the years 2001, 2006 and 2011.[2] The results from the wage decomposition show that there is a gradual shift away from migrants earning more than natives do, by £0.50 in 2001, to natives earning more than migrants, by £0.48 by 2011. This change implies that there has been a change in the overall wage distribution in the UK over this period.

[2] At the end of the thesis there are also tables for the full eleven year period.

I also find that there is a difference in the time it takes for male workers and female workers, who are natives, to match the average individual from the migrant groups earnings and then to exceed them. Due to this, I ran two more wage decompositions. One of these was to decompose male wages and the other was to decompose female wages.[3] I also attempt to show that this may have more to do with the type of job that individuals from the migrant group have a tendency to be employed in rather than any outright discrimination. The tables for the two different genders show that native males surpass the equivalent migrant by 2006. Also I show that native women take longer to match and then surpass the average wage for migrant women. This does not occur until 2008, two years after native males have overtaken the equivalent group of migrant male workers.[4]

The third chapter also attempts to explain why there is a fundamental change in the way that the wages are distributed between native and migrant workers. In order to study this question we must first find a way in which to examine any discrimination that may exist, towards migrants, in the UK labour market. We must first find a way in which to divide a representative population into two groups. One of these groups consists of UK native workers, the baseline category. The other group contains migrant workers who have migrated to the UK between 2001 and 2011. This is studied by using an Oaxaca-Blinder wage decomposition technique that allowed for the investigation of any potential discrimination in the UK's labour market directed primarily towards the group of migrants.

This chapter seeks to explain any gap that happens to exist in the average hourly wage rate that is received by native and migrant workers in the UK labour market. If any gap does exist in the earnings between the two groups further reasons as to why this may arise need to be examined. Is this due to one set of workers receiving higher wages due to having a higher level of skills, and is therefore better at the job, or if one set is rewarded at a higher rate just because they are a "preferred" group, and could therefore be classed as discrimination. In order to do this I shall be employing the work of Oaxaca (1973) and Blinder (1973) and using their technique of wage decomposition to

[3] These decompositions, as all others in this thesis is between native and migrants, be they a whole population or the male and female populations.
[4] This is shown in Tables 3.15 for men over the full ten year period and Table 3.16 for the female population.

investigate if there is negative discrimination occurring in the British labour market[5] against the migrant workforce. The variables in the Mincer-type wage equations are education, experience, the square of experience, tenure and dummy variables to show if an individual was a migrants, married or a female. It could also be the case that the migrant cohort before the two EU accessions was composed of individuals that just had a higher level of skills. My look into the occupations that the two different migrant populations are employed in will tell the reader that the migration of Eastern Europeans from the accession states has changed the type of jobs that migrants are tending towards being employed in.

The estimation of the decomposition for the two groups showed that two of the variables were the most prominent in determining the amount of wages paid to either of the cohorts. The first of these variables is the amount of education that an individual has received. The second of the variables that has a large effect on wages is not experience on its own. It is the experience that an individual has accrued in their current job, the variable for which is tenure.

Even with this, there is still a distinct difference between where the two groups find work. As is shown in the appendix in Table 3.2 there is a marked increase in the number of migrants being employed in "elementary" occupations, with an increase from 16.9% to 23.1%. This is also combined with a fall in the proportion of migrants employed in jobs as Managers & Directors (from 11.6% in 2001 to 4% in 2011).

This chapter also shows how the enlargement of the EU, to include the new nations that joined in 2004 and 2007, has changed the wages that are being earned by the two groups. I have gathered evidence that shows the type of work that these new migrants are being employed in compared to the native population. In order to get a fair grasp of some of the changes that the accession, and the UK's reaction to it, have been brought upon the UK labour market.

In addition, I investigate a set of counterfactuals for the two populations that are being studied. The counterfactuals that are presented in this chapter are taking the two populations endowments and applying the returns that were estimated for the other

[5] Discrimination in the context of this chapter may refer to one group just being payed less than the other because they are a member of that group or that group may have characteristics that face a penalty in employment opportunities.

group. For instance the returns that have been estimated for the native population are applied to the endowments that the migrant population have and from this we are able to calculate an estimated counterfactual wage that would represent an alternative mean wage if migrants were rewarded in the same way as the native population. This is then also applied to the subpopulations of the two genders when they are studied separately. As the endowments that the migrant group tends to be larger than the native group the estimated counterfactuals present some interesting results. The estimated counterfactuals show, at least up until 2006, that there should be a wage gap that is even larger than that which has been estimated between the two groups. I show in this chapter that if migrants were rewarded the same way as natives they would have been earning 1.88 log points (£6.55) in 2001 and by 2011 it would have risen to 2.19 log points (£8.94). This also happens for the two individual genders also both follow very similar patterns.

Chapter Four then presents the conclusion to this thesis. This conclusion will also point to avenues of further research on the topic of migration of foreign-born individuals now residing in the UK.

Chapter Two

Return Migration from the Host Nation

2.1 Introduction

Migration to and from the UK has always a major topic, both politically and economically. The coming and going of migrant populations has become even more important sine the expansion of the European Union. A more important aspect of migration is what type of individual is the average type of migrant from each country of origin that arrives in the UK, and what are the characteristics that these individuals who have migrated to the host country and have then, after a period of time, returned to their country of origin[6] have. This is a very important part of the story behind migration. For instance, a country may be experiencing a labour shortage in its low tech manufacturing industry. If this was the case then the average migrant that the government should be trying to attract are those from the lower end of the skill distribution in the source countries, the negatively selected. However, it may actually be attracting the opposite type of migrant, those from the top of the source countries distribution of skills.

Then what needs investigation is, given the initial flow of migrants, what is the average type of migrant that returns home versus the average type of migrants who becomes a permanent resident?

This chapter investigates the type of selection that characterises those migrants who have migrated out of the UK,[7] to return to their country of origin also known as the return migrants.[8]

The first thing to do in an investigation such as this, and the primary focus of this chapter, is to investigate the factors that attract migrants to the UK and what factors may incline those migrants to decide to return to their country of origin.

To do this a grouped probit model will be applied to a dataset that has been constructed from various sources. This could potentially be important for government policy on

[6] In this paper source country is where a migrant originated from and the host country is referring to the country where the migrant has migrated to, in the context of this paper the host nation is the UK, in particular England and Wales.
[7] In this paper the assumption is made that if a migrant from country *j* is no longer counted by the Census as being in the UK then they must have returned back to their source country.

[8] For the purposes of this paper, as in many of the other papers on this topic, return migration, remigration and outward migration are all considered to be synonymous with each other.

migration to the UK as individuals resident in England and Wales make up the vast majority of the UK population, 88.8%.[9]

This chapter aims to add to the analyses done by many different authors, but does so with data for the UK. The investigation conducted in this chapter has not been done for UK for this particular period of time. This is of importance for the UK as over the time span that is being investigated the government policy on migration experienced two exogenous shocks, primarily from the accession of the EU8 and EU2 to the European Union and the UK governments decision not to impose transitional restrictions upon the accession the EU8 countries, on 01/05/2004 for the EU8.

This chapter will also make the contribution of investigating the marginal effects of the grouped probit model. The marginal effects were not investigated by Borjas and Bratsberg (1996) or by any other paper that is examining the effects of the outmigration of individual migrants to a host nation.

2.2 Literature Review

2.2.1 "Who Returns? The Outmigration of the Foreign Born." (1996)

This work is primarily based upon Borjas and Bratsberg (1996). In this paper the authors investigated the type of selection that characterises those migrants who become permanent residents in the US. The authors argue that the traditional way to think about migration, that those who migrate to the US are the highest skilled individuals from the source nation and examine the selection process that determines the skills of the population who have migrated and the subset of this population who have chosen to reside in the US on a permanent basis. Individuals may choose to leave the host country as their initial decision was based on a lack of information that has led them to "fail" or bad luck may have led to the same result, which results in the individual returning to their source country.

In Borjas and Bratsberg (1996) the authors provide a mathematical formulation of the Roy model that was originally presented in Roy (1951).[10] The authors choose to look into migration flows into and out of the US in the period between 1970 and 1980

[9] http://www.ons.gov.uk/ons/guide-method/compendiums/compendium-of-uk-statistics/population-and-migration/index.html
[10] The model was originally presented in Borjas (1987) but was extended in Borjas and Bratsberg (1996) to include return migration.

and make the assumption that those not counted in the 1980 census have migrated back to their country of origin rather than onto a third country.[11]

Three different models are used to see whether those who remain, the permanent migrants, are either the "best of the best" or "the worst of the worst". The first of these models is a grouped probit model[12] which examines how in and outmigration rates are affected by macroeconomic factors changing in the host and source countries.

An augmented Mincer equation is then presented by Borjas and Bratsberg (1996) which shows the relative incomes of the various groups[13] who have migrated to and are now working in the USA. The final model is designed to show that as more migrants who are from countries that mainly provide migrants who are positively (negatively) selected arrive in the host nation then wages decrease (increase) for those from the same group who remain. The reason for this is that as more migrants arrive from positively selected countries then the skills of those who arrived before them are being diluted by the new arrivals and vice versa for those who can be characterised as being negatively selected. The two types of selection that can be generated through the movements of various groups of migrants is shown in Figures 2.1 and 2.2 in the appendix to this chapter.

In this chapter it will primarily be the first economic model that I am replicating.[14] This is done using Stata's command for grouped probit.[15] I am using this model to investigate how the migration rate changes given changes in macroeconomic variables.[16] The results from Borjas and Bratsberg (1996) first model have been replicated in Appendix Table A1.

The augmented Mincer equation is designed to examine the average relative wage that is received by individuals from the various source countries. This relative wage is then used in the third and final model to see how changes in the migration of different

[11] This assumption will also be used in this paper as well however the date that individuals are no longer in this country and are therefore assumed to be return migrants is the Census date in 2011 which is 27/03/2011 and is only for England and Wales rather than the whole of the UK.

[12] A Grouped Probit model works much like an individual probit model however the data is organised as blocks rather than as individual observations. The results should be the same as if it was done in a normal probit model.

[13] Groups and countries of origin are for the purposes of the paper synonymous.

[14] Table 2 from Borjas and Bratsberg (1996) is replicated in Table A1 in the Appendix.

[15] The commands for grouped probit analysis in stata are bprobit, with generates results through Maximum Likelihood, and gprobit, that estimates the coefficients that are to be estimated through Weighted Least Squares (WLS).

[16] It is assumed that all individuals are equally affected by changes in their source countries aggregate factors equally.

groups have an effect on the wages earned as estimated by the augmented Mincer equation.

From this third model the authors study how changes in the migration flows into the US may change the average wage that is received by that particular group of migrants. This is making the assumption that if migrants come from a country that has, in the past, provided highly skilled migrants then those who follow will tend to be the same as those who arrived in previous periods. This final model then allows the authors to draw the conclusion that those migrants who remain are either the very best or the very worst individuals in terms of skills.

One of the major drawbacks of Borjas and Bratsberg (1996) is that in their paper they only present the coefficients from their grouped probit. However probit coefficients, by themselves, only specifies the direction of any effect. In order to fully investigate the migration flows into and out of any country using a probit model I must look into the marginal effects of the model. The marginal effects are be presented in the results section in Table 2.2.

2.2.2 Selection Papers

Roy (1951) provides the seminal work on the process that determines the selection that would be observed given a number of assumptions about some hypothetical society. In this society there are only two occupations, hunting and fishing. Hunting is much easier than fishing, but has a much lower spread in the potential output. As fishing is harder, relative to hunting, and can potentially lead to a much larger output it is likely that those who fish are going to be better off than those who hunt. However how the two skill sets mix with individuals in society is another major concern. If skill in hunting and fishing are positively correlated with each other, then those who are good hunters are likely to fish given the potential rewards that fishing can bring. Even if they are negatively correlated then those who are only just below average at both tasks are likely to take up fishing as an occupation due to the potential rewards being higher for fishing than for hunting. This pattern of behaviour remains even after the society is extended to have more occupations. The main thrust is that the variance of potential output gets higher with the increase in skill level required.

Willis and Rosen (1979) examine the self-selection that occurs between those that do attend college and those that do not. The authors hypothesise that the decision is based

upon an individual's evaluation of their potential earnings given their choice of whether to attend college or not. The authors find that there is positive selection bias in the results of their probit estimation in both groupings of actual outcomes. The results would indicate that, on average, the correct selection was made, those that choose B over A earned more than if they had chosen A instead, even with identical characteristics. The downside of this paper, which is admitted to by the authors themselves, is that the data may well come from a source, military records, that would give somewhat biased characteristics, and that the observations were only included if they had provided the information on their first recorded income.

Selection in the context of membership of trade unions is examined in Lee (1978). In his paper he finds that those who join trade unions are of a higher skilled group than those who do not. Miller (1984) also finds a similar pattern but just for employment. Those who survive in the job longer, have a longer tenure, tend to be those who are best suited for the job. This may be due to the individual gaining some form of enjoyment from the job or having been weeded out by the company as not being productive enough, relative to what the company's initial expectations for this particular employee.

Many of the papers in the migration literature also look into the selection process that decides on the type of individual who does end up as being a migrant or even for those that choose to return to their country of origin, those who are return migrants.

Chiswick (1999) finds that the type of selection that characterises the migrants that he was studying tended to be decided by the costs that would be incurred during and after migration. He found that the average migrant would tend to be positively selected. This is due to not just the fact that those individuals would be better able to afford the monetary cost of any migration, but would also be better at adapting to the new surroundings that they were presented with after the actual migration (this would include, but is not limited to, being in a better position in which to learn a new language). These findings are much less pronounced in political and tied migrants as they had much less choice in where migration would leave them and why they were migrating.[17]

[17] Political migrants migrate on the basis of changing political regimes in the source country that may affect them, either religiously or politically. Tied migrants generally migrate as another family member has migrated and they have chosen to migrate along with that individual.

Later in the paper Chiswick gives a formal presentation of a model that shows how this situation may arise naturally. Even when the opportunities are equal in both countries for those of differing ability, in this paper the author only defines migrants as being high or low skilled, it will be the case that most migrants will be positively selected. The reason for this is that even if the direct costs associated with migration do not vary given the individuals level of ability there will be an advantage for those with higher ability in terms of efficiency of migrating. This may also, but not necessarily mean, that those who are more efficient in the migration process will earn more than those who not as efficient, although this would be clear.[18] One thing that does become apparent is that migration cost can never be considered as being the same for every individual from a source country. Even if the direct out of pocket costs are identical the opportunity cost will be larger for those with a higher level of skills. This is the case as the earnings that are forgone will be larger for those that are in the high skilled category than for those in the category of those with a lower level of skills. The main drawback, identified by the author himself, is that this paper does not work with the observed outcomes for any individual migrant or group of migrants.

Then there are papers that investigate the selection of those individuals who do choose to migrate. An early paper that looked into both migration and selection was Borjas (1987). Like Borjas and Bratsberg (1996), Borjas (1987) uses a mathematical formalisation of the Roy model to investigate the self-selection that determines the level of skills that the average migrant brings from their country of origin.

In this paper the author is only investigating one direction of travel of migrants and not the possible return or migrations onto third countries. The author wishes to show that the selection of the migrants by skill level is determined by the economic variables in the source country relative to those same variables in the US. The author claims that the level of economic inequality in earnings, relative to that observed in the US, will determines the average type of migrant from a specific source country that will arrive in the US. For migrants to be viewed as being positively selected, high skilled, one condition needs to be met. The wage structure in the origin country has to be less unequal in the source country than it is in the US. This will mean that those with high levels of skills could do better in the US than they could do at home. For migrants to be

[18] Those who are more efficient in migration are those with higher levels of skills than those who are not as efficient.

negatively selected the opposite needs to be true, it must be the case that income is more unequally divided at home then it is in the US. If this is the case migrating is used as a form of insurance given the level of skills of the migrant and we observe a migration flow that is negatively selected.

Jasso and Rosenzweig (1990) give a formal critique of Borjas (1987). In this paper they criticise Borjas' use of income accruing to the top 10% over that accruing to the bottom 20% as being the main way to observe whether a migration flow is positively or negatively selected. They also criticise the selection of countries which are selected as the origins of migrant flows. They continue to claim that Borjas has used an unseen selection criteria when deciding on which countries to include and which to exclude.

Another paper that uses the model developed by Borjas and Bratsberg (1996) is that done by Co, Yun and Gang (1998). This confirms the predictions of the original paper in their examination of migration, both into and out of the Hungarian economy. The authors use the Hungarian Household Panel Survey (HHPS) drawn from the panels of 1993 and 1994, as these panels were the first to ask if the individuals surveyed had any experience of working abroad. They also restrict their sample to only those of working age when they were surveyed. After accounting for the self-selection the authors find that men have a 4% wage premium and women a wage premium of 45% for those who have worked abroad in comparison to those who stayed. They also find that for those who work in wholly Hungarian owned firms suffer from a wage penalty of 30% and 27% for men and women respectively. These results are obtained through estimation by maximum likelihood. They also report estimation by Ordinary Least Squares (OLS). Generally the estimation by OLS gives some quite different results to that when the estimation is run with Maximum Likelihood. The Maximum Likelihood Estimation (MLE) gives a fairly significant difference of the earnings of those who have gone and worked abroad. For men the estimation gives a result that is significantly lower than that when the estimation is done with OLS. For women the opposite tends to occur. The overestimation of male earnings, after they arrive back from being abroad is to be expected when using OLS and MLE. However, for women there is a jump in the relative earnings for those who have worked abroad versus the women who have never left Hungary in order to work. The authors state that this could be due to unobservable characteristics of the female return migrant population. They state that these characteristics may make the women less desirable in the Hungarian Labour market

when compared to the labour markets in foreign countries.[19] They, however, do not look into what these attributes are and the value of them.

Rooth and Saarela (2006) attempt to investigate the selection that results from the migration between Finland and Sweden. As Finland has a much higher return to formal education than Sweden (9% per year in Finland versus 4% per year in Sweden) those who migrate from Finland to Sweden are the migrants that are generally characterised as being negatively selected. The theory given in Borjas and Bratsberg (1996) is again confirmed by the author's investigation of those migrants who return back to Sweden by 1995. The return migration results in an accentuation of the selection process. Those that end up returning can be considered as the "best of the worst" of those that were closest to being the marginal migrants.

Another paper that is primarily based on Borjas and Bratsberg (1996) is that done by Chiquiar and Hanson (2005). In this paper the authors aim to investigate the effect that return migrants have on the source country after their return. They also examine whether the results that are to be expected hold true. Borjas and Bratsberg (1996) state that if a source nation has a tighter wage distribution that is more centred around its mean and has a higher rate of return to skills, then the migrants will, on average, be more likely to be positively selected. This should imply that migrants to the USA from Mexico should be negatively selected, due primarily to the levels of income inequality. The migrants from Mexico should therefore be negatively selected.

However once the wage distributions are examined for Mexican natives in Mexico and those who have migrated, the authors conclude that the migration from Mexico to the US has been mostly from the middle to upper middle part of the distribution. This would then imply that those who do migrate are not simply negatively or positively selected but somewhere in between. Upon their return to Mexico these return migrants would normally fall into the upper or the upper middle parts of the migrant earnings distribution

Barrett and O'Connell (2000) set out to find the outcomes faced by Irish migrants if they choose to return home after time spent abroad. The author follows Co, Gang and Yun (1998) in order to test to see if migrants from Ireland in the 1990's receive a wage

[19] The authors do not mention the exact foreign countries that the migrants tend to be heading to. They only mention, and have dummy variables that account for, if the individual has gone to OECD countries in Europe, OECD countries outwith of Europe and non-OECD countries.

premium upon their return. The authors identify a number of waves of migration that occurred in Ireland during the late 20th century. They identify those who graduated from third level colleges in 1992 as being the most suitable group to examine. The beginning of this research tries to identify those who tend to migrate. The author supposes that, much like in the 1980's the migrants will tend to be highly educated.

The results published by the authors, after correcting for gender, show that men make a significant premium for migrating out of Ireland and then returning. However women earn about the same as if they had never migrated.

The authors also propose some alternative explanations as to why the return migrants tend to earn more than those who never migrated. One of their explanations as to why there might be a significant wage gap between the two populations is that the very fact that these people have migrated could be interpreted by employers that these individuals have better unobserved characteristics that would be desirable in the domestic labour market. In the paper they also report, from their own survey, why many people chose to migrate. By far the largest proportion of individuals who migrated said that they did so for "adventure" rather than for better labour market outcomes. There is potential bias in their results looking into the selection of the migrants as they are only looking into individuals who graduated from college or university meaning that the individuals are more likely to be positively selected.

In their paper De Coulon and Piracha (2004) attempt to use the model developed by Borjas and Bratsberg to investigate how emigration affects the source country. The authors apply the model developed by Borjas and Bratsberg to persons who have migrated abroad and then have chosen to return to Albania, however when investigating the selection that characterises migrants they apply the econometric models developed by Lee (1978, 1982). This approach allows the authors to be able to calculate the mean conditional wages of those who stayed in Albania and those migrants who left but then returned at a later date. This is especially useful in an Albanian context as much of the population have engaged in some form of return migration. They also use a semi parametric approached that was developed by DiNardo et al (1996) in order to estimate the change that migration has had on the wage distribution of Albania. They do this in order to examine the whole wage distribution for both return migrants and those who stayed in Albania. It is concluded that many of the migrants who derive from Albania

were negatively selected. The authors then posit the idea that if the stayers had chosen to migrate they would be earning much more. The reason, as proposed by De Coulon and Piracha, is that if the higher skilled did migrate then they would face much higher costs in their migration but would have also receive a higher payoff upon returning to Albania. Then they would, in all likelihood, have received some form of training on parts of the job that would require foreign language skills whereas those with a lower level of skills, and therefore working in more routine jobs, may not be required to have as much training in the host countries language.

There are also a number of other papers that discuss migration, both internationally and within states, but do not use the Roy model in order to investigate the effects of selection.

Interprovincial migration in Canada is studied by Robinson and Tomes (1982). They find that the expected outcomes from migration, moving from low income states to high income states, is generally supported by their findings. Also highlighted in the paper are the effects of native language on migration outcomes. Those who only speak French are less likely to engage in migration than those who are bilingual or English speakers.

Zhoa (2002) examines migration within China, mostly rural to urban and then the return trips made by the original cohorts of migrants. A big reason for this study was the response of cities in China to large migration flows into their area. The response was designed to deter migrants from arriving in these cities or to convince those that had that returning may be their best option. However the paper also shows that many of the migrants did not return home but migrated onto a third area. The main reason for migration, as hypothesised by the author, is primarily the lack of non-farm economic activity that can be undertaken by individuals within the town, the choice in effect falls to that between migration and the non-farm economic activity that there is.

Also, after their return, migrants are more likely to invest in farming equipment than are those individuals who have never migrated. This is primarily due to the fact that those who have migrated will, if they do, return with savings.

2.3 Model

The model that is to be applied in this paper is an adaptation of the Borjas Selection model that was initially presented in *Self-Selection and the Earnings of Immigrants* (1987). The model was adapted by Borjas and Bratsberg (1996) in order to incorporate for the flow of return migrants from the host country back to the source country. Three equations are required in order to describe the log wage distributions for the three different groups.[20]

$$w_0 = \mu_0 + \eta v \quad (1)$$

$$w_1 = \mu_1 + v + \varepsilon \quad (2)$$

$$w_{10} = \pi w_1 + (1 - \pi)(w_0 + \kappa) \quad (3)$$

For clarity, I assume that w_0 refers to the log wage earned in the source country where an individual originates from and w_1 gives the same but for the migrants in the host country, which for the remainder of this chapter is the United Kingdom. When we are describing the log wages faced by those who are engaging in return migration, I refer to w_{10} that represents the log wage distribution in both countries faced by those who do return. In these equations μ_0 describes the mean wage in the source country and in the host country the mean wage, μ_1, would be observed, if all individuals from the source nation migrated to the host country. The symbols v and ε are both deviations from the mean income. In this way v can be thought of as transferable skills, across borders, however ε has a better interpretation as misinformation in the migration decision or the luck of any particular individual. From these two facts we can state that the parameter v is known to anyone considering migration but ε remains unknown to all individuals. The other part in equation (1) that needs to be defined is η which is simply the rate of return to skills in the source country relative to the host. The final equation, equation (3), the log wage distribution for those individuals who first engage in migration from the source to the host country and then after a fixed period of time, described by the parameter π, upon return to their source country. If this does happen for any individual, it is assumed that there is a reward for spending part of the working life in another country. This reward is defined as κ and is to be thought of as a percentage above μ_0. At this point two additional assumptions must be introduced. The first of these is that for an initial migration to occur this would be at the beginning of said individuals working life. The next is that, as mentioned earlier, the length of time spent in the host

[20] The three groups are non-migrants and return migrants but also one-way migrants, those who move abroad but do not return.

country by an individual, π, is constant. This then implies that any individual engaging in the initial migration decision is doing so at the age they were at the start of their working life plus π years.

These migration decisions are reached once any individual has incorporated in actual costs that would be involved in the migration or the migration and remigration decisions. In the previous papers by Borjas[21] these migration costs were calculated as time equivalent costs. This means the easiest way to define the initial cost of migration is as $M=C_m/w_0$ and in terms of the costs of return migration as $R=C_r/w_0$. In these two cost functions C_m and C_r are the actually costs, in terms of pound sterling, that the individual would have to pay in order to arrive at the required destination. We can now state what conditions must apply in order for any individual to migrate or given that they have initially migrated to then engage in return migration. The initial migration decision would occur if:

$$\max[\mathrm{E}w_1 - M, Ew_{10} - M - R] > w_0 \quad (4)$$

Return migration would then occur if the above were true with the extra condition:

$$\max[Ew_0 - R, Ew_{10} - R] > w_1 \quad (5)$$

Equation (4) states that an individual chooses to migrate if the expected wage in the host country minus any migration costs exceeds what is currently available to them in the source country. For those who, after period π has come to an end, choose to return to the host country then the wage upon return minus the costs of both migrations must exceed what is currently available in the source country. Equation (5) then states the conditions that must arise for any individual to engage in return migration and incorporates the two possible time frames for return migration. If the individual "fails" in the host country, they have drawn from the distribution of g(ε) a significantly bad outcome, then they immediately return to the source nation. However if they get an outcome, that can be thought of as beneficial to the migrant, from their draw from the distribution g(ε) then these individuals may engage in return migration if the expected wage upon return exceeds what they were making in the host country minus any costs incurred due to the decision to remigrate.

[21] Borjas (1987) & Borjas & Bratsberg (1996)

For an initial migration decision to be changed into a return migration decision, in order for some investment to be made, we must add an additional assumption. This additional assumption will then assume that the returns to re-migration, κ, must exceed any costs that are incurred by moving between countries.

$$\kappa > M + \frac{R}{1-\pi} \tag{6}$$

Equations (4) and (5) can then be rearranged so that we can apply some sorting in equilibrium conditions. If an individual does decide not to migrate then it must be the case that:

$$(1-\eta)v \leq (\mu_0 - \mu_1 + \kappa) + \frac{M+R-\kappa}{\pi} \tag{7}$$

For all those who wish to migrate to the host country then it must be that the opposite of the (7) is true such that:

$$(1-\eta)v > (\mu_0 - \mu_1 + \kappa) + \frac{M+R-\kappa}{\pi} \tag{8}$$

Then we need a third equation to describe the behaviour of those who end up conducting return migration:

$$(\mu_0 - \mu_1 + \kappa) + \frac{M+R-\kappa}{\pi} < (1-\eta)v < (\mu_0 - \mu_1 + \kappa) + \frac{R}{1-\pi} - \varepsilon \tag{9}$$

It has now become obvious that the critical value that describes the selection that characterises migrants from a certain country is η. The value that η takes describes the average type of migrant that arrives in the host nation from any given source country. If η takes a value that is less than 1 then the average migrant could be described as being positively selected. This means that such an individual has been drawn from those individuals that are more than the average skill level when we examine the source countries distribution of skills. If this is the case then the migrants who return to the source country will be the least skilled of this current sub sample. They return as conditions at home, the source country, improve or if w_{10} ends up being greater for them than w_1 currently is. This will mainly be derived from the improvement in their wage in the source country granted by time earning a wage in the host country, and is represented by κ. Then it follows that those who only engage in one time migration tend to be those who are drawn from the very top of the sub-sample that did migrate to the host nation, if the flow from that particular nation is positively selected.

The opposite is true if η>1. If this is the case then the initial migrants will generally be characterised as being from the lower end of the source countries skill distribution, they can be classed as being negatively selected. Again it is those closest to the mean in the source country that could be considered as marginal migrants, those most likely to remigrate due to an improvement in the income they would earn upon return, κ, or due to a change in the situation in the source or host countries.

If both of these selection procedures are true then the return migration is acting to further the process of selection over the initial group of migrants in the host country. The authors of the original paper on this pointed to the fact that the further selection that occurs due to outmigration means that the host nation will be left with workers who are "the best of the best" or "the worst of the worst."

The above equations allow us to formulate the return migration probability function. This function is simply equation (8) and equation (9) formed into a fraction with (8) as the numerator and (9) as the denominator.

$$q = \frac{(1-\eta)v > (\mu_0 - \mu_1 + \kappa) + \frac{M+R-\kappa}{\pi}}{(\mu_0 - \mu_1 + \kappa) + \frac{M+R-\kappa}{\pi} < (1-\eta)v < (\mu_0 - \mu_1 + \kappa) + \frac{R}{1-\pi} - \varepsilon} \quad (10)$$

It is easy to show that for this equation, with no other assumptions, the derivatives for the main components are as follows:

$$\frac{\partial q}{\partial M} < 0, \frac{\partial q}{\partial R} < 0, \frac{\partial q}{\partial \eta} < 0, \frac{\partial q}{\partial \kappa} > 0 \quad (11)$$

These would imply that an increase in the costs of any direction of migration would decrease the amount of migrants travelling from any source country to the host nation. The wage premium upon return to the source nation would increase the outmigration rate from the host country as κ increases.

For this chapter, as in Borjas and Bratsberg (1996), I am making the assumption that those migrants who are no longer in the host country at time t' have returned to their source country. We are effectively excluding the possibility that the individuals have migrated onto a third country and are going to make the assumption that if an individual no longer appears in the records of the host country then they must have returned to their source country.

An assumption that is also required is that the populations of any source country are risk neutral. If it was assumed that they were risk averse, then, given the model that has just been derived nobody would want to migrate. This would be due to the migrants not wishing to engage in the random draw that would be part of determining their wages.

If on the other hand a countries population are to be considered as being risk lovers then it would be expected that the whole population of the source country would engage in migration. They would no longer be apprehensive of the result of the random draw from g(ε) as the migrant would be able to engage in return migration, or migrate onto a third country, if the draw from the distribution g(ε) was not actually a positive or a neutral result.

2.4 Data

The data required for this investigation, into the outmigration rate of foreign nationals from England and Wales, is extensive. For the data on the dependent variable, if individuals have returned back to their source country or not, we need to consult two separate sources of information. Data on the inflow from any single country was obtained from the DWP (Department for Work and Pensions), which can be broken down into data on the inflow from specific countries.[22] The implication of this, due to the lack of information on the individuals other than just the raw number of individuals who have been allocated a NiNo number, is that a grouped approach needs to be employed. The data gathered from the DWP represents the overall inflow of individuals, from foreign countries, in the time period t.

The other data source required for the dependent variables are those who are left in the country by the census date in 2011 and who arrived after the census date in 2001 and this is obtained from the 2011 UK census and represents those individuals that are still in the UK at time period t'. This was obtained from the Office of National Statistics (ONS).[23] This then allows for a calculation of those individuals who can be considered successes in term of probit, those who returned home.[24] This then allows the dependent

[22] NiNo allocated to foreign can be obtained through Stat-X on https://www.gov.uk/government/collections/national-insurance-number-allocations-to-adult-overseas-nationals-entering-the-uk

[23] These numbers were obtained on request from the ONS in Table CT0111 which details the number of recorded individuals, by passport held, that are in the country on 27/03/11 and are between the ages of 25 to 64.

[24] The assumption is made that if the individual is no longer recorded as being in England or Wales they have returned to their source country.

variable to represent the percentage of those from any one country who have engaged in remigration.

Other data that needed to be obtained are for the macroeconomic variables and therefore require different sources of data that is shared by individuals within a group.[25] A variable that is not actually shared by all in the group is distance. However in this context we make the assumption that the distance between the capital of the source country and London is a close enough approximation. This variable measures the distance between the capital of the source country and London in term of thousands of miles and is used as a proxy for the cost of migration.

Table 2.1 – Descriptive Statistics for Various Nations

Variable	Number of Observations	Mean	Standard Deviation	Minimum	Maximum
Inflow I(t)	99	32866.02	69047.02	120	567860
Census Final Stock R(t')	99	18247.57	42360.04	22	361373
Outmigration Rate q(t, t')	99	1.760738	3.56549	0.1486486	28.841176
Log GDP Gap	99	-1.409975	1.225613	-4.60378	0.5351179
Distance (000's of miles)	99	3.348384	2.288119	0.2	11.688
Income Inequality	99	17.91414	17.16927	3.3	87.2
Political Stability	97	60.95773	22.23949	17.2	100.5

Note: t represents the period of time when migrants are arriving in the UK. t' represents the point in time at the end of the analysis to represents the number of migrants that are still in the UK.q(t,t') represents the outmigration rate for each individual group.

Another of these variables is the gap in the level of GDP between the source country and the GDP of the United Kingdom. This is calculated by dividing the level of the GDP in the source country by the level of GDP in the UK. The data for this variable was obtained from UN estimates[26] of GDP for 2011.

Relative income inequality is also a factor that may well cause movements in population in ways that are described by the model. The information for this variable was obtained from the UN Human Development Report published in 2009.[27] This report gives details of the national income in any country and the percentage that is accrued by the top 10%

[25] In this paper the groups represent different nationalities.

[26] http://data.un.org/Data.aspx?q=GDP+per+capita&d=SNAAMA&f=grID%3a101%3bcurrID%3aUSD%3bpcFlag%3a1

[27] http://hdr.undp.org/sites/default/files/reports/269/hdr_2009_en_complete.pdf pages 195-198

and that which accrues to the bottom 10%. These numbers are then used, in the same way as in Borjas & Bratsberg (1996), but they use the income being made by the bottom 20%. These numbers are then placed into a fraction, with the income accruing to the top as the numerator and that being made by those at the bottom as the denominator. This then provides a single value that is used to represent the inequality in the distribution of income, which is theorised to have an effect on the rate at which migrants decide to return home. Even though this variable does not change, as a fixed point in time is used, it should be able to be used as a representation of the gap in the inequality of income.

The political stability of a country is also a major factor in an individual's decision about the country in which he or she is willing to reside. The data for this variable was obtained from The Failed States Index of 2011.[28] This index provides a number for the perceived stability of a nation's political structure. The score for any single country is the addition of scores for many different factors. This includes demographic pressures, human flight, legitimacy of the state and many others, However in this study the score for human flight has been removed as it includes what is actually being investigated by this current study.

2.5 Methodology

Borjas and Bratsberg (1996) investigate the selection of migrants who move from the USA back to their countries of origin.[29] In their paper they study how a number of macroeconomic variables may have an effect on an individual's decision to migrate for a second time in order to return home[30] by t'. I used the framework presented here in this paper in order to examine why migrants who come to the UK may choose to return to their source country.

In order to study why these individuals decided to leave the host nation a probit model was run on the data described in section 2.4 of this work. The aim of this chapter is to investigate the variables that have the largest effect on rates of remigration. This chapter will explore how the same factors, as used in Borjas and Bratsberg's paper, may affect

[28] http://ffp.statesindex.org/rankings-2011-sortable

[29]Borjas & Bratsberg (1996), *Who Leaves? The Outmigration of the Foreign Born*, The Review of Economics and Statistics, Vol. 78(1), 165-176.

[30] The economic variables of interest are the source countries GDP, its political stability, distance between the source country and the closest port to the source country in the USA and the country of origins level of income inequality.

migrants in England and Wales, and how changes in these variables may change their decision of being resident in the UK due to the changing situations in their source and the host countries.[31]

$$y_i^* = x_i'\beta + \varepsilon_i, \qquad y_i = 1[y_i^* > 0] \tag{12}$$

Equation (12) is the index function for the probit model. As y_i^* is not observed in the data an index function is required. If y_i^* is less than or equal to 0 then we set y_i as being equal to 0. If y_i^* is greater than 0 then we assume that y_i is equal to 1. All of the variables that will be used to examine the effect on the dependent variable are contained in the vector x_i, while the error term for the model is ε_i. The probit model uses the standard normal distribution and therefor the variance of the error term is equal to 1. In this model observations of the variables that are assumed to have an effect on the probability of whether an individual will decide to return to their source country and are included to explain why individuals choose to return home or not.

A number of other assumptions also need to be satisfied if I am to run a probit model. The error tem, ε_i, is assumed to be equal to zero. The error term is homoscedastic in the model. Another assumption is that each observation is independent of all on the other observations.

In a probit model the coefficients cannot be easily interpreted. Instead we need to look at the marginal effects that have been produced by the dependent variables by changes in the explanatory variables. I also take account of the various functional forms that any equation being estimated contains.

$$P(y_i = 1|x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + \beta_4 \log(x_3) \tag{13}$$

Equation (13) gives a simplified version of the model that is present in Borjas and Bratsberg (1996) which contains the 3 functional forms that I have encountered in the grouped probit model that is ran in this paper. Those functional forms are linear in x_1, quadratic in x_2 and logarithmic in x_3. In order to obtain the marginal effects the partial differentials must be taken for each term, and are shown below.

The marginal effect for x_1 is below in equation (14).

[31] It is assumed that if someone is no longer counted as being in England or Wales as recorded by the national census in 2011 has returned back to the source country.

$$\frac{\delta P(y=1|x)}{\delta x_1} = \phi(x\beta)\beta_1 \quad (14)$$

In equation (14) we are presented with the partial derivative for the coefficient of x_1. In the equation above $\phi(.)$ represents the pdf of the standard normal distribution. As can be seen in equation (14) this is substantially different from the standard OLS coefficient that would be associated with x_1 which would be just β_1. This is primarily due to the fact that the whole functional form is being multiplied by the pdf of the standard normal distribution.

$$\frac{\delta P(y=1|x)}{\delta x_2} = \phi(x\beta)(\beta_2 + 2\beta_3 x_2) \quad (15)$$

Equation (15) above shows the marginal effect for the quadratic term that appears in equation (13) or any other quadratic term that may appear in a model that is to be estimated in a probit model.

$$\frac{\delta P(y=1|x)}{\delta \log(x_3)} = \phi(x\beta)(\frac{1}{x_3}) \quad (16)$$

The above equation, equation (16) gives the marginal effect for any logarithmic variable, such as x_3 in equation (13).

One of the major advantages of using marginal effects is that one may examine the marginal effect at different points in the distribution. In this paper I will be examining the variables at their means, when $x=\bar{x}$. This is also the case when we are examining the individual groups. When this is done I am not looking at the overall average for the whole data set but at the mean for a particular variable in that group.

The number of groups in the dataset is:

$$n_g = \#g \quad (17)$$

This equation is stating that each observation, n_{gi}, where each i is part of a group, g.

$$y_{gi} = X'_{gi}\beta + \varepsilon_{gi} \quad (18)$$

Equation (18) gives the model that will be estimated by grouped probit where individuals within a group are indexed by i and the different groups are indexed by g. Each individual within a group has the same covariates as the other individuals, as the variables in $X_{gi}=X_g$ as they are macroeconomic variables and therefore do not vary

across individuals within the same group. This means that the equation that will be estimated is slightly different from equation (18).

$$y_g = X_g'\beta + \varepsilon_g \quad (19)$$

This causes issues in the estimation process, which will need to be resolved. In this paper I employ two of these in the estimation process. The first that will be employed is using errors that are robust to the effects of clustering and then weighted least squares is also be used. Clustering will not affect the values of the coefficients, however it will affect the standard errors of the model that is being estimated. We can apply an estimate of the variance that does not rely on the assumption that the errors in a group, or cluster, are independent.

$$\hat{V}_C = (X'X)^{-1}(\sum_{j=1}^{M} \varphi_g{}'\varphi_g)(X'X)^{-1} \quad (20)$$

The variance is estimated with equation (20).[32] This equation allows for correlation between individuals within a group, but still maintains the assumption that individuals in different groups are uncorrelated with each other. In equation (20) φ_j is the within cluster weighted sums of the contributions to $\delta y_{gi}/\delta\beta_{gi}$ and allows for the correlation between individuals within the same group j. φ_j is equal to $\Sigma_{j=1}^{M} X_j{}'u_j$.

The major drawback of using cluster robust variance estimation is that it can be inconsistent when the number of clusters is small.[33] However the number of clusters is considered to be small when it is about 30 or 40 clusters. This paper has 93 groups each of which can be considered a cluster.

The other method is to apply weighted least squares (WLS hereafter). This is required as OLS will be inefficient in the estimation of equation (19) due to the heteroscedasticity within the model. In order to do this each observation is weighted by some other variable. Stata allows weighting by the raw number of observations within a group, both the "successes" and the "failures." The variance for this is then estimated by equation (21).

$$\sigma_g^2 = \frac{p_g(1-p_g)}{n_g\phi^2\{\Phi^{-1}(p_g)\}} \quad (21)$$

[32] file://homefs03/home/vfj1/cluster%20robust.pdf – page 10

[33] Wooldridge, J, M *Cluster-Sample Methods in Applied Econometrics*, The American Economic Review, Vol. 93(2), 133-138

In equation (21)[34] we have p_g, the percentage of "successes", n_g, the overall numbers of individuals within a group, ϕ_g is the standard normal cumulative distribution function and Φ is the probability density function of the standard normal distribution.

Once equation (21) has been estimated then the model may be estimated with $1/\sigma_j^2$ being used to weight the different groups. This should, like with the cluster robust estimate of the variance, allow for a more efficient estimation of the coefficients.

2.6 Results

I show the results from the regression of log wages on a number of macroeconomic factors in Table 2.1. The first of these methods is via approximating the variance using cluster robust variance estimation. This was discussed in the methodology (2.5) section of this chapter. Another way that the clustering of errors can be overcome is to employ weighted least squares. This was also discussed in the methodology section of this chapter. When Table 2.1 is examined, it can be seen that the cluster robust estimation is done in column (2) of Table 2.1 for the outmigration rate and column (5) for the immigration rate. The columns of the table that used weighted least squares are column (3) for the out-migration rate and column (6) for the rate of immigration.

As can be seen in the results many of the estimated coefficients have a high degree of volatility in their estimated standard errors, in the various estimations, that is being caused by the correlation of the errors for the individuals in one group. This is primarily due to the issue of all the observations of individuals, indexed by *i* in equation (18), all sharing the same variable values in each country of origin that is indexed by the subscript *g* and hence driving a clustering of the error term. The different ways to correct for this are cluster robust variance estimation and weighted least squares, both bring the t statistics, and there for the standard errors, to a much more plausible level.

These results only point to the direction of travel and I had estimate the marginal effects for these to be able to represent a number of individuals that are affected by such a change. As was mentioned in the methodology section I do this through taking the partial derivatives of the various estimated coefficients. This is easily done through

[34] http://www.stata.com/manuals13/rglogit.pdf

most statistical packages.[35] The results for the marginal effects are presented in Table 2.2.

In Table 2.2 we can see how the various factors would affect the number of migrants engaging in return migration if any of these variables increases. Much like in the previous table, Table 2.1, the differing specifications give similar numbers for those wishing to return home, given an increase in any of the individual factors that are theorised to affect the flow of migrants, however the issue is again the standard errors assigned to these factors when it is being estimated by the standard grouped probit.

It can be seen from the estimations that are presented in Table 1.2 that the estimation with the most consistent standard errors are presented in those columns that estimate the standard errors using cluster robust estimation of the variance or weighted least squares. It can be seen from these columns that the most important factors regarding the estimated variables are distance and the gap in GDP between the two countries.

Distance, which as mentioned before is a proxy for the cost of migration, has an estimated value of -0.182. The table shows the reader that as the cost of the migration increases there will be fewer individuals willing to engage in migrating to the host country.

The gap in the logs of GDP also has a high level of significance in its estimation. The value that was estimated in the grouped probit estimation was 0.0781. This would indicate that as the gap in GDP between the host and source country increases there are more individuals wanting to move to the UK.

There are also the estimates of how income inequality can have an effect on the outward migration of those who have already migrated to the UK. It can be seen from Table 1.2 that the value that has been estimated in column (2) is only 0.0495. If this is compared to the same variable, but for the rate of immigration, it can be seen that this is much larger than it is for the rate of return migration to the various source countries with the coefficient that is estimated for this being equal to 0.0772.

We can now look into the estimated numbers for the immigration rate and how the variables have an effect on it. Those who have migrated to the UK and plan to go back

[35] The command for marginal effects in Stata is “margins” and the command that was used in this study to examine the marginal effect at the mean, by group, was “margins, dydx(*) atmeans over (ccode)”.

are less likely to return to a poor country, one that has had a coup or a revolution in the recent past and one that will cost them more to travel to. Most of the estimates from this regression have a negative sign before the estimated number. This shows that as the values for these variables increase individuals are less likely to want to return.

All of the other variables in Table 2.1 are significant, but only at the lower threshold of 25% level of significance as derived from their standard errors. This implies that the estimates of the other variables need to be used in helping to explain the changes in migration rates but are still not representing the full story.

The estimated values for the rate at which individuals immigrate are both, relatively, larger and more precisely estimated. If we stay focused on the WLS estimation it can be seen that the estimated value for income inequality on the immigration rate is 0.0272. This can be compared with the score that this variable receives for the outmigration rate which are much lower than those estimated for the rate of return migration.

The gap in the level of GDP, as measured in 2011, does provide evidence that an increase in the GDP of the source country would lead to more migrants returning to their home country. This is confirmed in the marginal effects of an increase in the gap between host and source countries level of GDP. Marginal effects, given the weighted least squares specification, would indicate that an increase in the level of GDP in the source country would lead 926 more individuals deciding to return to their source country. Looking into the two source countries that are specified in Table 1.4 would indicate that if GDP increased in Poland then there would be over 14000 more Poles willing to return home, whereas only 7661 more Indians would return due to an increase in the GDP of India.[36]

The stability of the source countries political regime is also a major factor when individuals are making decisions about their migration strategies. It is theorised by Borjas and Bratsberg that a more stable political system in the home nation may result in more individuals returning to their home country or just less people leaving the source country. The coefficients estimated by the model would indicate that this is correct, even though the t stat for it would indicate that it is only significant at the 25%

[36] GDP per capita in Poland in 2011 was $20200 while the GDP per capita in India was $3680.

level which could easily imply that the true value for the coefficient for this is actually zero.

The marginal effects for this variable indicates that 42 more people would choose to migrate back home if the score for political stability increases. Once the marginal effect, at the mean is examined for Poland and India we get a better idea of the effects of political stability on the migration rate. If we examine a change in the political stability score, that would indicate that 662 Poles would choose to return while if India's score for political stability increased then by the same amount then 350 more Indians would choose to migrate back to India. This becomes clear as the score for political stability increases as nation become more politically unstable so any increase in the score of political stability can be viewed, in reality, as an increase in political instability and therefor an increase in the score for this variable would be more likely to induce less people to return to that source country.

The last variable to be examined is Income Inequality. An increase in the unequal distribution of income may work both ways. The theory would indicate that an individual's decision to migrate based on the level of income inequality is dependent on the average selection of the individual migrant from country g. The score that is given by the model, no matter its specification, would indicate that income inequality is not a factor that individuals would base the intentions to migrate based on.[37] However once the marginal effects are investigated we get a better picture of the actual effect. The marginal effect given by the model would indicate that 57 less individuals would choose to migrate if the level of income inequality at home increases.

The two source countries have quite similar levels of income inequality: India's score is 7.3 while Poland's is 8.8.[38] Even though the levels of income inequality are much the same we find, once we look into the marginal effects, that the amount of people from each are almost double for Poland than it is for India. In fact, an increase in the level of income inequality in India would lead to 765 less individuals returning, while an increase in the same variable in Poland would lead to almost twice as many people engaging in return migration.

[37] The coefficient, in each specification, is zero to the third decimal place.

[38] The mean score for Income Inequality in the data set is 17.48.

This chapter has shown that two variables are more important than others in determining migration flows, to and from the country. The first of these variables is the cost of migration, which in this chapter has had distance. Another important variable, that helps to determine the size of flows from certain countries, is the gap in GDP between the two nations.

2.7 Conclusion

This chapter applies the model developed by Borjas and Bratsberg (1996) to data from the UK for the period from 29/04/01 until 27/04/11. The analysis required the construction of a new data set. This required combining a diverse range of other data sources including information from the 2011 UK census and information derived from the UK government's Department for Work and Pensions. These two data sources were used to calculate the rate of return migration for the 93 countries that were studied. The construction of this data set also allows the potential to look into the average selection that decides the type of migrant that comes from different countries.

I find that the two biggest factors that can affect outmigration rates, for different groups, are the distance between the source capital and London and the gap between the source and host countries GDP in 2011. I also calculate the marginal effects of each variable which has not been done before.

The marginal effects of an increase in distance was found to be 692[39] more individuals returning home if the source was 1000 miles further away from the UK. There was a larger reaction to the narrowing of the gap in the two countries GDP. It was found that if the gap closed by 1% there would be an additional 1091 individuals returning to the source country.

I also estimated the marginal effects of the change in variables for the two largest migrant cohorts in the UK over the time that has been selected for study. These countries are Poland and India. These two countries contradict many of the points from the initial paper by Borjas and Bratsberg (1996). Distance from London, which was selected as a proxy for travel costs, to New Dehli is over 5,000 miles more than it is to Warsaw, and is therefore considered more expensive to travel to and from. Almost

[39] The precise number estimated was 691.57 which was found through the specification of grouped probit through maximum likelihood.

twice as many people are going to travel between the capital of India (10,904) and London than would between the London and Warsaw (5,760).[40] This would point to the opposite of the observation that was in the original paper, where the further away the source country is the fewer people will return.

It could be, however, that in the 21st century and the decreases in transportation costs that have accompanied this, many of the migrants from countries close to the UK, mostly EU nations, may find it easier to return home for social and family events. While this is happening for the average Pole, the distance for the average Indian is much further, implying higher costs, and they may wait until they remigrate before engaging in social situations.

Once we investigate the two countries differing reactions to a decrease in the gap between the GDP of host and source country it can be seen that similar reaction is being experienced by migrants of India and Poland. If there is a decrease in the GDP gap between India and the UK this is estimated to induce 17,204 more individuals to return to India. The same situation happening between the UK and Poland is estimated to lead to 9,087 more Poles returning back to their home country.

[40] Numbers in parenthesis are the estimated numbers returning to country of origin.

Chapter Three

Is There Discrimination Against Migrants in the UK Labour Market?

An Oaxaca-Blinder Wage Decomposition

3.1 Introduction

An investigation into gaps in the level of wages that are earned by different groups can give an indication that a labour market is exhibiting signs of discrimination against a particular group of individuals. I use an Oaxaca-Blinder wage decomposition in order to identify any gaps in the wages earned between workers who are native to the UK, and those who have migrated to the country. This will allow me to be able to separate those aspects that have a direct effect on wages and that part which can only be labelled as unexplained.

In this chapter, I will provide a short literature review on the economics of discrimination and then review papers concerning the Oaxaca-Blinder wage decomposition. This technique allows me to identify reasons why there are differences in the average rate of pay that the two groups are remunerated for the labour that they provide. After differences in the levels of endowments have been accounted for I will be left with a gap that is unexplainable, which can also be interpreted as discrimination.

Using this technique allows me to examine the gap in pay between UK native workers and migrant workers in the UK over the study period, 2001 until 2011, and will allow for an analysis in the changing nature of the pay gap between these two groups. This is pertinent in the context of the UK, as over this period there was two shocks to the stock of workers in the UK's labour market. I find that the accession of the EU8 and the EU2 had a significant effect upon where migrant workers fitted into the wage distribution in the UK.

3.2 Literature Review

This literature review is broken into two parts. The first part will look into a number of papers that describe discrimination within the labour market. It will then take an extensive look into the literature on the wage decomposition theory that was developed by Oaxaca (1973) and Blinder (1973) as the basis for assessing the differences between the mean wages that are received by the two groups.

3.2.1 A Short Review of the Literature on Discrimination

Becker (1957) produced a book which examines the interaction between the discipline of economics and the practice of discrimination, based on one group being over or undervalued when it is compared to another group. It could be that members of one of

the groups do not want to work with any members of the other group. This will normally not be based on evidence about the discriminated against individuals of the other group.

Becker proposed that white workers would have to be paid a premium relative to the black workforce in order to attract more white workers into the organisation, leading to a gap in wages and lower pay for the minority group. This does not require that both the employer and employee to discriminate against the group of black workers, only that one of the two have preference for not working with or employing this group or to have a "taste for discrimination" as Becker puts it in his book. He goes on to introduce a discrimination coefficient that is added to a wage rate, price charged by firms too and availability of products or services to the group that is discriminated against (in favour of) by employers, sellers and other economic interactions that these two groups have with each other, including co-workers of the non-discriminated against group.

It could also be the case that employers are discriminating against one of the groups, and therefore discriminating in favour of the other. This makes the assumption that there are only two groups of workers. This is exactly what Becker does in his book by specifying that there are only white and black workers in the market.

Becker indicates that there are a number of different reasons why one of the groups is being discriminated against. The first reason has to do with the sheer numbers in the minority group of workers, with larger numbers leading to an increase in discrimination. The other point has to do with the dispersion of the minority group in the host country.

Becker theorised that the majority group will require an addition to their wage if they are being made to work with members of the minority group. It could also be the case that employers do not see workers from the minority group as having the same level of skills that the majority group has, this is not based on observation only a belief by those who are employing them. If it were the case that the members of the minority group had lower marginal productivities than the other group then employers would be justified paying them at a lower wage rate than members of the majority group, who are more productive.

Arrow (1998) discusses racial discrimination in a more modern context. Even with the passing of the US Civil Rights Act (1964) it can still be considered that discrimination

based upon the colour of an individual's skin still exists. This can be due to a number of reasons.

It could be that individuals with black skin are discriminated against due to the size and influence of their social networks. Arrow notes that an individual's beliefs and preferences, in a free society, are not governed by prices and markets. In a market subject to competitive pressures, discrimination should disappear.

Due to the fact that social interactions are not expressed in a market atmosphere social discrimination may well persist. He argues that even with markets an individual's beliefs will play a role in any transaction and that discrimination may well persist.

3.2.2 A Selective Review of the Wage Decomposition Literature

The literature on wage decomposition is fairly extensive. There are two papers that led to the introduction of the Oaxaca-Blinder wage decomposition technique and from where the technique's name is derived.

Oaxaca (1973) was one of the first to try to explain wage differentials between two well defined groups. This paper aims to estimate the gap in earnings between male a female full time wage earners in USA. It uses data from 1967 Survey of Economic Opportunity in order to estimate wage gaps between male and female workers. Oaxaca found that there is no difference in pay among most types of jobs. The discrimination that can be observed is not that women were not paid the same as men for doing the same job, but rather that the type of work that they were involved in was quite different for the average female compared with the average man.

The other central paper in this line of investigation is Blinder (1973). This paper centres more on wage disparities between the black and the white male populations in the USA as well as male and female populations and uses data from Michigan Survey Research Centre's "Panel Study of Income Dynamics".

Blinder begins by examining the wage differential between white and black males. The author's estimations show that whites earn, on average, 50.8% more than the black population. Most of this difference is due to the white population having larger returns to endowments. On the other hand, the black sample gains more than the white by being a member of a trade union, seasonal employment and from tenure on the job.

The paper also presents a decomposition of the male-female wage differential. The author provides the reader with an equation for the following wage decomposition. The female wage is estimated, by the author, to be 45.8% of the male wage. The largest single variable that favours males is estimated to be the return to age. Blinder points out that females have an almost flat wage-age profile while males display the standard convex profile as their age increases. Men also display larger coefficients than women for their return to education and dealing with local labour market conditions. The results show a very similar effect for the variable for local labour market conditions.

Reimers (1983) seeks to investigate the earnings differential that is experience by US men of Hispanic or black origin. The author initially lists the wage differential of various Hispanic groups that are present in the US. The author does point to the fact that this could be a language issue. This is because the fluency rate of English is lower amongst all the categories of non-white Hispanic individuals than it is amongst the white population and the population of immigrants in general. The author also points to the possibility that this may not be aimed at Hispanics in general but against all groups of migrants who cannot speak the English language fluently. He also highlights that many immigrants, especially those from South America, have lower rates of human capital than the native population has. Both of these issues can be circumvented. The state could provide more education to the groups that are most affected by the disparity in educational attainment and help with assimilating migrants quicker than they currently are. This could be done, in part, by giving immigrants lessons in English lessons which should minimize the integration issues that many face upon arrival.

In Neumark (1988) the author offers an alternative to the standard Oaxaca-Blinder wage decomposition technique. The author proposes that given the absence of discrimination (both positive and negative) in the labour market that the wages for both genders would be somewhere in between the current observed wages for males and that for females. However this non-discriminatory wage does not exist in reality and therefor needs to be estimated. He also shows how this model would also apply if there was discrimination that only worked in one direction for one of the groups[41] and therefore one of the existing wage structures can be thought of as not exhibiting discrimination.

[41] Much like in a lot of wage decomposition literature the examples used by the author are male/female differentials, but this could as easily apply to differentials by race or other personal characteristics.

Cotton (1988) disagrees with both Oaxaca (1973) and Blinder (1973). He argues that the techniques developed by both authors fail to regard that for a single population, composed of various groups, discrimination can work in both directions. It may be the case that one group within the population is more favoured by the labour market. The author then sets out a slightly different way in which the wage decomposition can be done. If a non-discriminatory set of coefficients can be estimated then the wage decomposition can proceed as normal. However, when the difference in coefficients is to be estimated the process is slightly more complicated. Instead of the estimation of differentials in coefficients just being those coefficients estimated for the majority group minus that same coefficient for the minority group as we now have a non-discriminatory coefficient.[42] This changes the calculation of the wage decomposition. Firstly, the non-discriminatory wage has to be taken away from the majority group, and then the minority coefficient is subtracted from the non-discriminatory coefficient. In the paper the author shows the benefit of being a majority worker combined with the cost of being a minority worker.

Differing occupational distributions are, at least partly, responsible for wage inequality between two groups, and this is investigated by Brown et al. (1980). In this paper the authors attempt to correct for any bias that may be introduced into the Oaxaca-Blinder wage decomposition arising from differences in the distribution of occupation between the two genders. In order to alleviate the bias that this may introduce the authors run a multinomial logit model on the male population and then apply the results to the female population. Differences in the proportions in work in each population and the hypothetical distribution that would result from women being employed on the same basis as men are added to the Oaxaca-Blinder decomposition. The results that are found lead the authors to suggest that most of the gender gap in wages may be due to the occupational difference that are observed between the two genders. They find that if women were employed in a similar way to the way that men are employed there would be many more females in managerial positions and less doing more clerical occupations.

Miller (1987) attempts to replicate the technique used in Brown et al., but this paper focuses on the UK rather than the US. His paper follows the same process as that of

[42] This coefficient should be somewhere in between those coefficients for the majority and minority groups.

Brown et al. in that the first stage is to estimate a counterfactual distribution for females across occupations, but Miller uses a multinomial probit model to estimate the probability of how many women in the sample would be in different jobs compared to the current distribution. As in Brown et al., Miller found that it was not inter-occupational differences that drove the discrepancy in pay between the two genders. The differential in pay he found was more the cause of intra-occupational difference in pay rates. The unjustifiable difference in pay rates that is not down solely to differences in socio economic variable between the two groups is more down to the fact that women are less likely to be promoted in their company than a comparable man is.

Polavieja (2004) investigates, as others have also done, how the type of job that is generally occupied by females may be a factor that is driving the wage gap between the genders. He proposes that the division in distributions of "task specificity"[43] jobs are that the higher degree of task specificity that is required by certain jobs will generally lead to that occupation receiving higher wages than jobs that only have low specificity. It is argued by the author that the inclusion of a variable that defines the specificity of the tasks involved in a certain job could be a useful addition to the Oaxaca-Blinder wage decomposition methodology.

Some jobs may be friendlier to those who may take extended periods of time out of the labour market.[44] The author checks this against a model developed by Goldthorpe (2001). In this model two jobs are proposed. Job A has initially a lower wage than job B, but this is quickly exceeded by the return to tenure related pay increases due to the specificity of the job. Job B, as mentioned above, starts with a higher rate of pay than job A. However job B has a much flatter curve when describing the returns to experience.

This point was explored in the early work of Mincer and Polachek (1974). In this paper the authors aim to explain reasons why women, on average, earn less than a male over their working life. This paper proposes that female earnings are less than an equivalent man due to females engaging in more time spent not participating in the labour market. The authors propose three different life cycles engaged in by the average working

[43] Job specificity is referred to in the paper by Polavieja and refers to how complex a job can be.

[44] The author specifies that this is more likely to be women due to potential time off due to child-rearing.

woman. If the female does not marry and does not have kids then women engaging in this form of working life should see a progression similar to that experienced by men. If the female does have kids then she, after the birth of the first child, may leave the labour force entirely and not return until after the children have gotten to school age. For those engaging in labour market participation of this kind Another possibility that may present itself is that the women has a break between child bearing and the children attending school. These individuals would generally be assumed to return to the work force after their children have started school. This would imply a double peak in earnings for anyone who does engage in this pattern of working over their life-cycle. The authors do admit that the period between child birth and the children going to school may be a period where the skills that the woman has accumulated over her life may depreciate.

Suh (2009) sets out to investigate, and decompose, the change in the inequality between male and female wages. The author runs an Oaxaca-Blinder decomposition on both 1989 and 2005 and then decomposes the changes over that period. The author uses the adaptation of the decomposition technique that was developed by Neumark (1988). The comparison between the 2 years show that in a number of the variables that would be counted as components for human capital the gap that was exhibited between males and females has narrowed or that women have even overtaken the average of those variables relative to men. This is especially noticeable in the length of time that the two groups spend in education. Even though the author has showed a narrowing of the gender wage gap he does concede that women, no matter which gender is used as reference, still earn roughly 80% of the equivalent male.

The difference between the earnings of natives and immigrants in Malaysia is examined by Annees et al. (2011). In Malaysia migrants into Malaysia have, on average, higher levels of Human Capital variables than native workers do.[45] Even with higher amounts of human capital, immigrant workers to Malaysia earn less than the average native Malaysian. Migrants will normally select themselves into a job in the manufacturing industry whereas the native wages tend more the select a job in the service sector. The average level of wages is higher in the services sector than it is in manufacturing

[45] One of the main examples that is referenced by the author is the length of time spent in education. The average length of time spent in education by a migrant is 15 years versus 13 for the average native worker.

industries than they are in the Malaysian manufacturing sector. They also highlight that there is a regional aspect to where the migrants have a tendency to be based. Differentials between regions could also explain some of the earnings differential between the two groups. The author then concludes that there is a degree of discrimination[46] leading to the difference in the average earnings between the two groups.

Bitilagy (2014) uses the Neumark-Oaxaca version of an Oaxaca-Blinder wage decomposition in order to attempt to explain the gender wage gap in Egypt over the late 1990's and early 2000's. Like many other studies, the author investigates if there is a gender wage gap in Egypt and, if there is, to attempt to explain why such a gap that the different genders are paid, on average exists.

One of the biggest problems that can occur when using the Oaxaca-Blinder technique is the identification issue that can occur when the regressions of the group specific wage equations contain dummy variables. Yun (2005) offers a potential solution to the problem. The author suggests estimating the wage equation using each possible value of the various dummy variables as the reference group. He demonstrates with an example using a set of dummy variables that will lead to varying coefficients for the different variables given that the reference group is changed. In the example in the paper he has three dummy variables. He shows the ramifications of using various categories as the reference point for the dummy variables and that the other variables coefficients are in reference to the variable that has been left out. He shows that the easiest way that this can be done is by estimating the wage equation using various categories as the reference variable. After this has been done, Yun suggests that for each category the easiest way to estimate the overall coefficient is to take the estimated coefficients that are generated and then take the average of these number of values generated.[47] In the simple example that he shows in the paper it is simply adding three numbers, one of which is zero, and taking the average.

3.3 EU expansion and wage discrimination between migrants and natives in the UK

[46] Defined by Oaxaca and Blinder as being the difference in the estimated intercept and estimated coefficients between the two groups.
[47] Obviously the one where this category is the reference group will be equal to 0

I now turn to the analysis of the differences between migrants and native wages in the UK using the wage decomposition techniques developed by Oaxaca (1973) and Blinder (1973). I will focus, in particular, on the period between the censuses of 2001 and 2011. This period is particularly interesting because of the large increase in migration to the UK as a result of the expansion of the EU and the extension of freedom of movement to workers from the new member states (NMS hereafter).

The EU8 are eight of the ten nations that joined the EU on 01/05/04 and since then have had free access to the UK labour and goods markets: Estonia, Latvia, Lithuania, Poland, Czech Republic, Slovakia, Slovenia and Hungary (the remaining two are Cyprus and Malta). The EU2 are Romania and Bulgaria who both joined the EU on 01/01/07. Since the accession of the EU8 and EU2 there has been a significant increase in migrant labour residing in the UK and therefore to the overall labour force. In 2004 the UK government was one of only three countries, along with Sweden and the Republic of Ireland, to allow free access to individuals residing within the new member states free access to their domestic labour markets (Longhi and Rokicka 2012). This however this was not repeated with the two states that joined in 2007.

The analysis in this chapter focuses on data drawn from three individual years – 2001, 2006 and 2011. The reason that these years were chosen is due to the accession of the new EU member states into the union. The reason that 2001 was chosen was due to the fact that this year predated any accession of the Eastern European states into the EU and therefore was able to provide a year showing the results of Ordinary Least Squares regression and the two different wage decompositions without any influence from the accession of the NMS. The reason 2006 was chosen was due to the fact that the EU8 had joined the EU on 1/5/04 and so had had 18 months to settle into the United Kingdom and find work. The final year that was chosen for this study was 2011. The reason for this is that all of the new member states had joined and the transitional arrangements had finished by January 2011, at least for the EU8. This meant that the EU8 and EU2 and the island of Malta and the whole island of Cyprus[48] were now members of the EU.

Table 3.1 shows descriptive statistics for natives and migrants, separately, in the UK for the years 2001, 2006 and 2011 based on the UK Labour Force Survey (LFS); the data source is discussed further below. Upon inspection of Table 3.1a, it can be seen

[48] Technically Cyprus is a divided state, however the whole island was allowed to join http://europa.eu/european-union/about-eu/countries/member-countries/cyprus_en

that migrants, on average, have higher levels of education and are more likely to be married than an average individual in the native population. However the native population, apart from in 2001, has a higher average wage than the group of migrants. Basic economic theory suggests that an increase in supply would normally lead to a decrease in the price of that good. This would suggest that a significant increase in the amount of migrant labour to a market would push down the wage in the occupations that migrants tend to get employment in.[49] This is discussed in Ruhs and Vargas-Silva (2015) in their report for the Migration Observatory, although the authors do highlight that this gap has more to do with the specific sectors that the migrants tend to be employed in and that even though it is significant the reduction in wages is not particularly large.

By running an Oaxaca-Blinder wage decomposition I can investigate how much of this wage differential is down to discrimination against migrants simply for being migrants and not being a native to the UK. This discrimination could simply be because they are migrants, or because migrants share some common and undesirable trait such as the inability to speak English as well as a native individual.

The Oaxaca-Blinder wage decomposition is widely used to investigate why one group of workers is paid more or less than another group. The technique is normally done as a two-fold decomposition, where the differences are cut down to those differences that can be explained by differences in the two groups endowments of the covariates that have a direct effect on the wages of an individual and unexplained portion, which is shown by the differences in the estimated coefficients. A threefold version of the wage decomposition also includes an interaction term that assesses both the gap in endowments multiplied by the gap in coefficients as these two items move together and not in isolation.

In this chapter's threefold decomposition analysis there is an extra step that looks to further decompose the part of the twofold decomposition that falls into the explained category. An inspection of the twofold and the threefold tables shows that the "Unexplained" section is exactly equal to the coefficients section of the threefold decomposition. The reason behind this can be seen from an examination of the equations for the two various decompositions. As can be seen from equation (7) there

[49] Dustmann et al (2005) gives an idea that wages in certain sectors may decline but also concludes that there will be benefits derived from incoming migrants.

is a bit more to the usual twofold decomposition as there is now the non-discriminated set of coefficients contained in the vector β*. In this analysis I have decided that the best non-discriminatory coefficients should be those for the native group. This changes the analysis slightly. The part of equation (7) on the far right hand side, everything within the square brackets changes once the analysis is done the way I have decided on using. By inspection it can be seen that the first part within the square brackets contains the term (βn-β*). However, by assumption, we know that those two terms are equal to each other and therefore that part of the equation is equal to 0. We now only have the term on the far right hand side, which we can see after a quick inspection of equation (9) is exactly the same. This allows me to further decompose the explained section of a twofold decomposition into that part that is down to the different endowments between the two groups and the interaction terms.

The analysis is done using the Stata statistical package and the Stata command "oaxaca" by Jann (2008). Details of this Stata command are presented in Jann (2008). As well as discussing the Stata command, the paper also discusses how the command produces either two-way or three-way decompositions. In a three-way decomposition, besides the endowment differential and the difference in coefficients between the two groups, there is also an interaction between the endowments and coefficients.

3.4 Data

The data for this study comes from the UK Labour Force Survey (LFS hereafter). The LFS is a survey aimed to be a representative survey of all private households within the UK. A wide range of data is collected for the LFS, including questions on earnings and the national origin of individuals. The LFS is a quarterly rotating panel survey of approximately 40,000 households.[50] Each household is interviewed in five consecutive quarters. Each household consists of one or more individuals, only some of whom may be working. The information collected makes it possible to compare earnings of native and migrant individual residents of the UK, taking into account individual characteristics such as age and education.

The three panels of Table 3.1 show the descriptive statistics for various dimensions of the populations under study. Table 3.1a shows the descriptive statistics for the overall population. Tables 3.1b and 3.1c have descriptive statistics for the subsets of men and

[50] LFS User Guide vol.1 – page 9

women. These tables show a number of interesting features. For the total sample, the average wage for migrants is initially well above that for natives in 2001. By 2006 the gap has almost disappeared, and by 2011 migrants have a substantially lower wage when they are compared with the native cohort. The patter is similar for men and women taken separately, except that native men have already overtaken migrant men by 2006, whereas native women overtake migrant women only by 2008.

It can be seen from these tables that the two populations have a number of differences in their endowments of human capital. The migrant population is younger than the native population. The various tables show that the normal age difference between the native and migrant population starts at about 2 years in 2001. The difference in the average age between the two groups grows throughout the sample period with the results shown in Table 1a showing and age difference of 2.13 years in 2001 compared with a gap of 5.67 years by 2011.

Another measure in which the two populations differ is the amount of years of education that have been completed. The migrant cohort always has a higher amount of education than the native cohort does; for example, 2.37 more years of education for women in 2001 to 1.85 years of difference in the 2001 male cohort.

It can also be seen from these tables that the average migrant has, on average, more than two years greater work experience than the equivalent native worker. This applies to both genders separately as well as for the sample as a whole.

Figure 3.1 – Wages of both Migrants and Natives

Source: Figure from http://www.migrationobservatory.ox.ac.uk/briefings/characteristics-and-outcomes-migrants-uk-labour-market, derived from ONS.

Table 3.2 shows the type of job that the migrant population ends up working in. By 2006 there is a sharp rise in the migrants population who work in the "Elementary" occupations. This carries on into the final year that is sampled, 2011, with a further rise in the proportion of the migrant population involved in "Elementary" occupations. Analysis by CEPS shows that most migrants, from the EU and beyond, are on average younger and better educated than the UK domestic population taken as a whole, as is the case in the LFS samples used in this study.

This study focuses on the hourly pay that each individual receives from working in their main job.[51] Respondents to the survey provide data on earnings in two of the five consecutive quarters, the first and the fifth quarters in which the individual is surveyed in. In this study I will only be using the responses from the earnings question in the first quarter in which any individual was surveyed. This is partly to address the potential bias that results from attrition; it could be that the household does not reply beyond a certain number of waves for various reasons.

In order to estimate the Mincer-type wage equations that are used for the wage decomposition, we require variables for wages, education, experience and an indicator of whether the individual is a native or a migrant to the UK. I also use indicators for gender and marital status. Several of these variables had to be derived from the data that is provided in the LFS.

A rate of pay that would be consistent across all individuals had to be derived. An important reason for this is that the LFS reports total earnings in a period irrespective of whether the person is employed full-time or part-time. It is also necessary because earnings in the survey are reported variously as yearly, monthly or weekly pay rates. Rates of hourly pay were derived by taking the total weekly pay that an individual is paid per week and dividing this number by the total number of hours that the same individual works in an average week.

The variable that is used for education requires a small computation in order to convert the given data into a form that is to be used in this study. The data that is reported by the LFS is simply when an individual leaves full time education. In order to derive the amount of years that an individual has spent in full time education I have taken the

[51] Earnings from secondary jobs were excluded for this study.

reported age at which they finished their full-time education and subtract 5 from this, as this is the age when most children will have started their schooling in the UK.[52]

Two different variables are used to cover the experience that an individual has accumulated in their working life. The first is the tenure that they have accumulated in their current job. The second is the amount of potential work experience over their working life before their current job. As is standard in the literature, the variable used to measure an individual's potential level of experience is their current age minus the age at which they completed their education. It is also standard to include in Mincer-type wage equations the square of experience as an explanatory variable. Experience-earnings profiles (or age-earnings profiles) typically have an inverse U-shaped curve

For the purposes of this chapter, anyone who reports in the LFS that they came from a foreign nation, i.e. a nation outside of the UK, is considered a "migrant" to the UK. A "native" is someone who reports themself to be British or to come from any of the individual nations that comprise the UK, England, Scotland, Wales and Northern Ireland, as being their nationality.

3.5 Methodology

This chapter will largely follow the methodology as described by Jann (2008). Firstly, we identify two different groups. Natives and migrants are the two groups of interest for this study. The task is to investigate the source of the difference in the dependent variable, the log of hourly wages. The log of hourly wages of individual *i* in group *g* is $\ln(w_{gi})$. We specify and estimate, separately, a wage equation for each group of the form

$$\ln(w_{gi}) = X'_{gi}\beta_g + \varepsilon_{gi} \quad (3)$$

The explanatory variables are X_{gi} with coefficients β_g and ε_{gi} is the error term.

As by assumption $E(\varepsilon_{gi})=0$ and β_g is a constant,

$$E(\ln(w_{gi})) = E(X_{gi})\beta_g \quad (4)$$

[52] https://www.citizensadvice.org.uk/education/school-education/access-to-education/#h-compulsory-school-age ; http://chartsbin.com/view/z87

We write the log of hourly wages of an individual in the native group as $\ln(w_{Ni})$, and similarly for an individual in the migrant group as $\ln(w_{Mi})$. Then the average difference, R, in the log hourly wages is

$$R = E(\ln(w_{Ni})) - E(\ln(w_{Mi})) \quad (5)$$

After the two regressions required have been estimated we can then use (5) to decompose the difference in the rates of pay, and its determinants, experienced by the two groups. The estimated difference in the log hourly wage rates is based on the difference in sample means:

$$\bar{R} = \overline{\ln(w_N)} - \overline{\ln(w_M)} \quad (5')$$

where $E(\ln(w_{Ni}))$ is estimated by $\overline{\ln(w_N)}$, the average log hourly wage for natives in the LFS data, and for migrants $E(\ln(w_{Mi}))$ is estimated by $\overline{\ln(w_M)}$.

Equation (4), for the two groups, is now substituted into equation (5), giving us equation (6). This enables us to decompose the differences in pay into differences between the two groups' average endowments of socioeconomic attributes (education, experience, gender, etc.) and into how a countries labour market rewards these two groups differently, or a coefficient effect. In both Oaxaca and Blinder the differences in the βs, i.e. the intercepts and the various coefficients under study, can be thought of as reflecting differences in the unobserved characteristics displayed by the two groups.

$$R = E(X'_{Ni})\beta_N - E(X'_{Mi})\beta_M \quad (6)$$

This can be estimated using sample averages for the endowments and the Ordinary Least Squares (OLS) estimates $\hat{\beta}$ for the coefficients:

$$\bar{R} = \bar{X}'_N\hat{\beta}_N - \bar{X}'_M\hat{\beta}_M \quad (6')$$

This shows the two contributing factors to the gap in wages displayed by the two groups. Equation (6) can be rearranged to decompose the wage gap into the contribution of endowments of socioeconomic variables and the returns to those variables. The wage decomposition can be calculated in one of two ways, both of which will be presented in this chapter.

The twofold decomposition is a good technique to use when a researcher suspects only one group suffers from discrimination. It could also be the case that one group is discriminated against in a negative fashion and the other receives no positive discrimination or that there is only positive discrimination and no one suffers from negative discrimination. This is appropriate in the application here, because it can be reasonable to treat natives as not facing any discrimination.

The twofold decomposition requires us to define β* which represents the coefficients that would hold if there was no discrimination in either a positive or a negative direction. Then we can rewrite equation (6) as:

$$R = \left(E(X'_{Ni}) - E(X'_{Mi})\right)\beta^* + [E(X'_{Ni})(\beta_N - \beta^*) + E(X'_{Mi})(\beta^* - \beta_M)] \quad (7)$$

Equation (7) shows the equation that is used in a twofold decomposition. The difference in the endowments between the two groups in the twofold decomposition is known as the "explained component".

As can be seen above the difference in the estimates in the gap in endowments are weighted be the hypothetical non-discriminatory wage equation coefficients included in the vector β^*. When we are looking into the gap in coefficients we need to weight them separately, using the individual group average endowments, depending on what part of the gap in coefficients that we are investigating. It is expected that in a labour market that has discrimination, or the over-valuing and/or under-valuing of unobserved characteristics, one of the groups, normally the majority group or in the context of this chapter the group of natives, will be paid more than the neutral group (with coefficients β^*) which would, in turn, be paid more than the minority group, or migrants in this chapter. It could also be the case that only one of the groups suffers from discrimination and the other group suffers from none. Here we will assume that the natives do not suffer the same kind of discrimination as do the group of migrants. We therefore set $\beta^* = \beta_N$ and obtain equation (8):

$$R = \left(E(X'_{Ni}) - E(X'_{Mi})\right)\beta_N + [E(X'_{Mi})(\beta_N - \beta_M)] \quad (8)$$

We are only interested in the difference between the migrant wage structure and that experienced by the group that suffers from no discrimination, which by assumption in this chapter is the group of natives.

It is also relatively simple to calculate a threefold wage decomposition. This is done by rewriting equation (6) as equation (9) below. The main difference between the twofold decompositions in equations (7) and (8) and the threefold decomposition in equation (9) is the fact that equation (9), as well as calculating the gap in endowments and coefficients also calculates the interaction between the gap in endowments and the gap in the estimated coefficients as they both move at the same time.

$$R = \big(E(X'_{Ni}) - E(X'_{Mi})\big)\beta_M + E(X'_{Ni})(\beta_N - \beta_M) + [E(X'_{Ni}) - E(X'_{Mi})](\beta_N - \beta_M) \quad (9)$$

Equation (9) allows the reader to investigate the different aspects of the threefold wage decomposition. The three parts on the right hand side of (9) are the difference in the two group's average level of endowments and the coefficients in the separate wage equations for the groups and the interactions between the gaps for average endowments multiplied by the gap between the two groups estimated coefficients. It makes sense that the coefficient effect part is included in equation (9) as that can be used to describe if there is any discrimination in the labour market, $E(X'_{Ni})(\beta_N - \beta_M)$. This part does not just include the difference in the estimated coefficients for the two groups but also the difference in the estimated intercepts that I estimated in the OLS regressions of the individual groups' wage equations.

If there are differences in the average observable productivity between the two groups then the reason for the difference is down to the vectors of X_N and X_M. In Becker (1962) it is argued that the development of human capital occurs throughout the working life of an individual. If UK society had no discrimination present in its labour market then the differences in the wage rates between the two groups would be down to the two groups' average level of socio-economic characteristics, their productivity. The differences in the marginal productivities of the two groups would explain why there are differences in the wage rates, on average, between the two. However, some employers may pay lower wages to migrants for various reasons: migrants may not have all of the skills that are required by the company, said employers may prefer natives for some reason or the employers may simply discriminate against migrants.

In the twofold decomposition, as in Tables 3.5, 3.8 and 3.11, the top panel gives the reader the average log wage received by the two groups for each of the years being

assessed. Below this appears the difference between the two groups and then the parts of that difference that fall into the categories of explained, the overall endowment effect, and the part of the gap in the log wage that is unexplained. It is gaps in the unexplained part that could be considered as discrimination.

Below this top panel there are two more panels. In these panels the individual variables are then assessed for their gaps in endowments and coefficients that allows for the second panel to be concerned about the explained gaps, gaps in the endowments, and then the third and final panel that breaks the unexplained part down for each individual variable. The intercept term is included in the unexplained part.

The threefold decomposition tables, Tables 3.6, 3.9 and 3.12, are structured in a very similar way to the twofold decomposition. In the top panel appears the average log hourly wages for the two groups, and the difference between these two averages is presented in the first three lines of these tables. After this, besides the breakdowns for the endowment and coefficient effect, now just called endowments and coefficients, there is a third component. This third part of the decomposition is labelled the interaction effect.

There are now four panels beneath this top section. Each of these describes the three effects for the individual variables. Much like in the twofold decomposition, these three panels break the overall effects into the effects of the individual variables.

3.6 Results

The results in this section are divided into three sections that are then each divided into four subsections. The first section will discuss the results for all migrants and all natives in the years 2001, 2006 and 2011. I will first discuss the results of the Ordinary Least Squares (OLS) estimates. I will then move onto the twofold decomposition for each of the two groups in each of the three years that are being studied. Then I discuss the corresponding results for the threefold decomposition. After the results for the decompositions have been reviewed, I then move onto a set of counterfactual estimates for the log per hourly wage. The counterfactual considered is for the migrant group to keep their initial endowments but to be rewarded for these endowments as the native population is and also the native population's endowments with the returns that the migrant group are estimated to have.

This is then repeated twice for the groups of subpopulations. The first of the subpopulations will be male-only groupings of natives and migrants. All four of the different processes will be performed again for the groupings of men. Then I consider the female subpopulations of the overall groups.

A complete set of twofold decompositions of all eleven years between 2001 and 2011 appears in the Appendix to this chapter.

3.6 Results

3.6.1 All migrants vs. all Natives

3.6.1.1 OLS results

The results from the estimation of the standard OLS wage equations are given in Table 3.4. As can be seen from a quick examination of the table, coefficient estimates for the group of natives are precisely estimated, and statistically significantly different from zero with p-values of 0.001 or smaller. The estimates for the group of migrants are less precisely determined, which is not surprising given the small numbers of migrants in the sample relative to the numbers of natives. The two variables for the migrant sample that are consistently different from zero at the 0.1% level are years spent in education and their tenure on the current job.

The return to education is consistently higher for natives than it is for migrants over the period being investigated. The estimated return to education falls slightly over the period for natives, but first increases and then decreases for migrants. By contrast, an individual's return to tenure in the current job in which they are employed in is higher for migrants than it is for natives, especially in the later years of the period being studied. A potential reason for this pattern could be due to the length of time since the initial migration, which is decreasing over the sample period.

3.6.1.2 Twofold Decomposition Results

Table 3.5 shows the results of the twofold Oaxaca-Blinder decomposition. The reported gaps are all weighted gaps in the endowments and the coefficients.[53] The mean hourly wage initially favours the migrant cohort that is present in 2001 by 8.3% (a negative difference indicates a gap in favour of the migrant group). These numbers change so

[53] This can be seen on page 54 of this work.

that by 2006 migrants earn approximately the same amount as natives; the log percentage difference is 1.4%, in favour of natives, but is not significantly different from zero. By 2011 it is the group of natives who are now earning a higher amount by 5.9% in log terms, on average, than the group of migrants.

We next look into the individual components of the decomposition. Table 3.5 shows that the 2001 difference in the explained, or endowment, part is equal to a gap of 0.14 or a log difference of 14% in favour of the migrant cohort. Most of the endowments of the variables in the explained portion contribute in favour of migrants; the exceptions are tenure in the current job,[54] married status and experience (which make contributions that are insignificantly different from zero, the latter when the effects of experience and it square are combined). The biggest gap, in favour of the migrant group is for education with a gap of 0.154, so that education is providing 15% difference in log terms to the gap in earnings between the native and migrant group and does so in the favour of the migrant cohort.

Once the unexplained, or coefficient, gap contributions for 2001 are examined it can be seen that the contributions are not very precisely determined and most components are not significantly different from zero. Only the contribution of the differences in the return to education is statistically significantly different from zero and only at the 5% level. This contribution is 0.322, so that natives have a higher return to education and this provides a wage advantage of just over 30%, but this is not precisely estimated with the standard error for this variable in 2001 being equal to 0.143.

The results for the 2006 decomposition are reported in the third column of Table 3.5. As has been mentioned above, the wage gap for 2006 is now slightly in favour of the native cohort, relative to the migrant cohort. The natives are now earning 1.958 log points, or £7.09, versus the migrant cohort's average earnings of 1.944 log points, or £6.99. Overall migrants are now earning slightly less on average than the native group by £0.10 per hour although this gap is not statistically significant.

The part of the decomposition that considered as explained by the difference in endowments has fallen in the period between 2001 and 2006 from -0.142 to -0.130 with both being in the favour of the migrant cohort. The unexplained part of the

[54] This contributes 2.6% in log terms towards the native wages vs. migrant wages

decomposition has also changed in the same direction as the explained part of the decomposition, but the change is much larger for the unexplained part than it is for the explained part. The change in the contribution to the native-migrant wage gap of changes in the endowments is 0.007 log points in favour of the natives (from -0.142 to -0.130), whereas the contribution of changes in coefficients is 0.084 log points (from +0.060 to +0.144).

If we examine the explained component it can be seen that the advantage that the migrant cohort has in education has decreased slightly by 0.009 log points, from -0.154 to -0.145. Much like in 2001, only the endowment of education in the explained section of the decomposition favours the group of migrants. The key characteristic in the explained section that favours the native group is tenure. When examining this variable and referring to Tables 3.1a it is clear that migrants are younger than the native population, and a large part in 2006 derive from Eastern Europe, the so-called EU8, countries that had only just joined the EU; therefore the majority of these migrants are new to the country. These new migrants can be expected on average to have fewer years in their current job than the cohort of migrants in place in 2001. Much like in the results of the decomposition for 2001, the total contribution of experience is not statistically different from zero, nor is marital status, but gender contributes about 0.01 log points in favour of migrants.

With respect to the contribution of the unexplained part of the decomposition one of the biggest variables that contributes in the favour of natives is the return to education. However, although the contribution of this variable is large (+0.0923 log points), it is imprecisely estimated: the standard error is 0.088, and the contribution is not significantly different from zero. By contrast the return to tenure is also large in absolute terms (-0.056 log points) and is estimated with more precision than the return to education, with the standard error for tenure in 2006 being -0.01.

The results for 2011 reflect a continuation of this trend with natives still having higher earnings relative to migrants. In 2011 the log average wages for natives and migrants are 2.097 and 2.037 log points respectively. The gap in wages earned by the two groups has increased to 0.0593 log points. This gap is estimated precisely, with a standard error of 0.016 and is statistically significantly different from zero at the 0.1% level.

In the decomposition for 2011 results for the explained section once again, as in 2001 and 2006, the migrants have a greater amount of education when they are surveyed. However, the total contribution of the education endowments to the wage gap in 2011 has not changed much from either the levels estimated in 2006 or 2001 (-0.152 log points in 2011, -0.149 in 2006 and -0.154 in 2001). The contribution of tenure in the job has moved slightly further in favour of natives, from +0.0332 in 2006 to +0.0439 in 2011. As in 2006, the contribution of endowments for both education and tenure are precisely estimated, with small standard errors. The contribution of tenure in 2011 to the unexplained portion is about the same as in 2006, at -0.0547 log points. The biggest change between 2006 and 2011 in terms of the unexplained effect is for the return to education. The 2011 estimated returns to education now favours the natives so that the contribution to the gap is 0.256 log points, compared to a contribution in 2006 that is not statistically significantly different from zero.

The complete set of annual twofold decompositions in the Appendix confirms these patterns. Between 2001 and 2011, the overall wage gap moves steadily from being in favour of migrants to being in favour of natives. The explained component – reflecting differences in the endowments of the two groups – is relatively steady over the full period, falling only slightly from about 13-15% in favour of migrants in the early part of the period, to about 11-12% by the end of the period. The unexplained component – reflecting differences in coefficients – starts in favour of natives by about 6-7% at the start of the period, and rises steadily to about 17-20% by the end of the period. Within the explained or endowments contribution, the educational endowment contribution in favour of migrants is steady throughout the period at about 15-17%, whereas the tenure endowment in favour of natives slowly increases from under 3% to around 4%. This reflects the influx of migrants; who are new arrivals and have spent less time in the country and hence in the host countries labour market. The patterns over time in the unexplained, or coefficient, contributions are also largely driven by education and tenure. The return to education in the early part of the period favours natives but is not precisely determined; by the middle of the period, it is insignificantly different from zero before a return to contributing in favour of native group. The return to tenure, by contrast, shows the interesting pattern discussed above. In the early part of the period, there is no significant contribution to the wage gap from differences in the returns to

tenure, but this changes steadily over the period to a substantial and precisely determined contribution in favour of migrants of about 4-6% by the end of the period.

3.6.1.3 Threefold Decomposition Results

The results of the threefold decomposition are presented in Table 3.6. As was covered in the methodology section the tables for the threefold decomposition are laid out slightly differently to the twofold decomposition. As with the twofold decomposition results, the gaps in endowments and coefficients are weighted gaps. The estimated wages follow the same pattern as they did in the twofold decomposition – the overall wages for natives and migrants and the gap between the two is the same by definition as in Table 3.5 for the twofold decomposition – but there is now an additional effect, the interaction effect, and the size of the endowment effect has altered. However due to the way that this is calculated the coefficient effect is identical to that estimated in the twofold decomposition.[55]

Much like in the twofold decompositions, the migrant cohort outperform the native group in terms of the endowment of education by 0.105 log points, or 10.5%, in 2001. This is repeated in 2006 and 2011 (the contributions of education are -0.131 in 2006 and -0.109 in 2011). All of the estimated gaps in education are significantly different from zero at the 0.1% level. By contrast, tenure in the current job in which an individual is employed favours natives in all three years. The size of this gap is 0.0302 in 2001, 0.0950 in 2006 and reaches 0.103 in 2011. These gaps are all positive and statistically significantly different from zero, and therefore favour the native cohort in the gap in hourly pay.

Next, we move on to the gap in the estimated coefficients for the two groups. Like in the twofold decomposition presented earlier, the number of gaps that are significantly different from no gap at all is less than those for the endowments between the two populations. Only four of the estimated coefficient gaps are significantly different from zero at the 5% level or less: two for the returns to education and two for the returns to tenure. The significant gaps for education appear in 2001 and then again in 2011. The first of these gaps, which is equal to 0.322 return to education in 2001, is precisely

[55] This can be seen on page 54 of this work.

estimated but is only significant at the 5% level. The second of these gaps is estimated with more precision at 0.256 and is significant at the higher 1% level. Both of these gaps, and the imprecisely estimated gap in the return to education in 2006, which is estimated to be 0.0923, favour the native population over the migrant one.

In addition to these gaps, the threefold decomposition, as discussed in the methodology section of this chapter, includes the effects of the gap in the interaction between the gaps in the native/migrant endowments of the various variables and the gaps in the estimated coefficients. In all three of the selected year's gaps in the interaction between coefficients and endowments favours the migrant population. The most statistically significant interaction gaps in this section mirror, to an extent, the most significant variables from the gap in the estimated coefficients. Rather unsurprisingly, it is once again the gap in the interaction for education in 2001 and 2011 and the gap in tenure in both 2006 and 2011 that are the only significant gaps in the threefold decomposition. All of the significant gaps in the interaction part of the decomposition reflect, even to the same levels of significance, the gaps that are produced in the coefficients section of the decomposition. The interaction between the gap in endowment and the gap in the estimated coefficients in education is equal to -0.0487 and is only statistically significant at the 5% level. The other gap which is statistically significant, at the 5% level, is the interaction for education for the 2011 estimate and is equal to -0.0427, which is significantly different from 0 at the 1% level. The other two gaps that are significant are for tenure in 2006 and 2011 that are estimated to be equal to -0.0618 in 2006 and -0.0593 for 2011.

3.6.1.4 Counterfactuals

Table 3.13 shows sets of counterfactual predictions of what the various populations would earn if they had the coefficients that have been estimated for the other group. For instance, if the migrant group had the same coefficients that were estimated for the native group, but the same level of endowments as they have as presented in Table 3.1.

The first row in Table 13 shows what the outcome would be if the group of migrants, overall, had the coefficients that are estimated for the group of natives in each of the three years being studied. By comparing the results from Table 3.13 to those found in Table 3.5, it can be seen that the overall group of migrants would be earning 1.88 log points (£6.55) per hour rather than the 1.841 log points (£6.30) that they were earning,

on average. This shows that in 2001 that migrants were being paid less than they would have been if they had the same returns as the native cohort does. This is continued throughout the period of study, with the prediction for the amount that would be earned by the migrant group being 2.07 (£7.92) in 2006 and 2.19 (£8.94) in 2011.

Counterfactuals were also calculated for the group of natives if they were rewarded as the group of migrants are for their level of skills. The counterfactual for the native group in 2001 is 1.77 log points (£5.87). This continues throughout the study period with an estimated counterfactual of 1.93 (£6.88) in 2006 and 2.04 (£7.69) for 2011.

The counterfactuals show an interesting story in this study. The average earnings for the migrant group would be substantially higher, apart from in 2001 when it is only marginally higher, if their endowments were rewarded on the same basis as those of natives.

3.6.2 Male Decompositions

In this section the results of the Ordinary Least Squares estimates and the twofold and threefold decompositions for male natives and migrants are presented in Tables 3.7, 3.8 and 3.9 respectively.

3.6.2.1 OLS Results

Table 3.7 presents the results of the OLS estimates for male natives and male migrants over three years in the sample period. For the male-only samples there are a number of small changes to the results for the wage equations for the native sample compared to the pooled male-female results that have been discussed above. For the migrant sample, only the returns to education and tenure are estimated with any precision and are significantly different from zero.

3.6.2.2 Twofold Decomposition

The top panel of Table 3.8 provides information on the average earnings for the two groups in the three selected years of study. It also provides the difference between the earnings of the two groups and how much of this difference is due to the endowment effect, the explained portion, and how much of this difference is due to the coefficient effect, labelled as unexplained. The next two panels are divided between the make-up of the explained difference, in the second panel, and the make-up of the unexplained

gap in wages in the third panel. The top panel shows that the insignificant gap between male migrants and natives in 2001 is the result of higher endowments of migrants (explained) offset by higher returns to these endowments of natives (unexplained). By 2006 and continuing into 2011, the pattern changes noticeably: the wage gap is in favour of natives by about 7-8% in both years, driven by an increase in the gap deriving from the coefficients (the contribution of the endowments themselves is essentially unchanged throughout the period under study, at about 10% in favour of the migrant group).

The explained part of the gap is presented in the second panel of the twofold decomposition. Only two of the differences in the initial endowments are statistically significant in all of the three years for which the decomposition has been estimated. The first of these is the gap for education that, on average, the two groups have. It can be seen from Table 3.8 that the contribution that the amount of education that the native group has is substantially less than that of the migrant group. Table 1b shows that over the three years for which results have been shown, the male migrant group has about two years more education than the equivalent native group and that the contribution of this gap in endowments shown in Table 3.8 is always significant at the 0.1% level. The other gap in endowments that is noteworthy is the gap in tenure. The contribution of tenure is always in favour of the male native group over the three years. This gap is increasing in each of the three years, starting off at 0.0428 log points in 2001 and rising to 0.0498 log points by 2011.

The bottom panel of Table 3.8 show the results for the twofold decompositions unexplained component, or the gap in the estimated coefficients. In the bottom panel, it can be seen that for the male group, differences in the return to education between the two groups do not contribute in a statistically significant way to the wage gap. Tenure, however, is important: the return to tenure for migrant males is significantly larger compared to that for native males in 2006 (-0.0845 log points) and 2011 and (-0.552 log points).

3.6.2.3 Threefold Decomposition

The full threefold decomposition for males can now be examined. The top panel again provides the average wage that is received by the two groups, the difference between these averages and the components that make the difference between the two average

wages and the gaps in the endowments, coefficients and the interaction between the two. In 2001, the insignificant gap between male migrants and male natives is the result of higher endowments that favour male migrants being offset by higher returns to these endowments for the native group. In 2006 and 2011 the pattern changes: a gap in favour of natives emerges, driven by higher returns to endowments and only partly offset by interaction effects.

The second panel allows a look into the weighted gaps of the initial endowments that the two groups have, on average. The key pattern here is the emergence of a gap attributable to job tenure: the tenure endowment contributions to the wage gap are 0.162 and 0.116 for 2006 and 2011. The two gaps in 2006 and 2011 are precisely determined, with standard errors of about 0.025 and 0.027. This gap in tenure makes sense: Table 3.3 shows that the majority of migrants in 2001 had arrived from Western Europe, but by 2011 this had changed to the countries that acceded to the EU in 2004 and 2007 respectively. These new workers were recent arrivals and would have only started to accumulate tenure after their arrival.

The third panel of Table 3.9 shows that although the endowments of tenure of newly arrived migrant males was lower than that of natives as just noted, the returns to tenure for these migrants was higher than for native males: -0.0845 log points in 2006 and -0.0552 log points in 2011. The final panel of Table 3.9 gives the details of the interaction terms. Again, the pattern for tenure is noteworthy and similar to that for the contribution of the job tenure coefficients, i.e. in favour of migrants: -0.123 log points in 2006 and -0.0660 in 2011.

3.6.2.4 Counterfactuals

The counterfactual estimates were repeated for the male-only grouping. This shows much the same picture as the overall group estimates of the counterfactual average earnings if the two groups were rewarded on the basis of the other groups' returns to the various endowments. Table 3.13 shows that in 2001, when the migrant endowments are combined with the native coefficients, the earnings, in order to provide the counterfactual, the differential between the two groups hardly changes: 1.83 log points (£6.23) vs. the actual log wage of 1.84 log points (£6.30). However by 2006 the migrant group would now earn 2.01 log points (£7.46) and in 2011 this rises to 2.15 log points (£8.58).

We now repeat the exercise using the endowments for the native group combined with the coefficients that are estimated for the migrant group. A very similar story emerges for the native counterfactual in 2001 as was given for the migrant counterfactual, namely almost no change: the counterfactual for male native earnings would be 1.77 log points (£5.87) vs. the actual 1.76 log points (£5.81). For later years, the native group takes an earnings penalty with the estimated counterfactuals being 1.84 log points (£6.30) earned by the native group of males in 2006 and 1.98 log points (£7.24) in 2011.

3.6.3 Female Decomposition

3.6.3.1 OLS Results

Table 3.10 provides the results for the wage equations, separately, for the female subpopulations of both natives and migrants. As can be seen from the table, all of the OLS coefficient estimates for the native group are precisely estimated and statistically significant at the 0.1% level. The coefficients for the migrant group are less precisely estimated, again as before due to the smaller sample size. The return to a migrant's education is always statistically significantly different from zero and is similar in magnitude to the estimated return to education for natives. The pattern for the return to tenure is also noteworthy. For natives, it is consistently about 1% across the period analysed. For migrants, in 2001 it is, statistically, no different from zero, but in 2006 and 2011 it is significant at the 0.1% level, with the levels estimated for 2006 and 2011 being equal to 0.0326 and 0.0236 respectfully, and so larger than the return to tenure for natives.

3.6.3.2 Twofold Decomposition

The results from the female-only twofold decomposition are presented in Table 3.11. As can be seen in the top panel the results show that, much like in the overall decomposition, migrant women earn 12% more than natives in 2001, a gap that is significantly different from zero. By 2006 this gap has disappeared – it is only about 2% and not significantly different from zero – and by 2011 it has reversed, with a statistically significant gap of about 6% in favour of natives.

The rest of the top panel splits these differences into explained and unexplained components. It shows that the explained component (endowments) is consistently in favour of migrant women, but the contribution is falling, from -0.151 in 2001 to -0.127

in 2006 and -0.107 in 2011. By contrast, the unexplained component (coefficients) starts in 2001 at only 0.0295 and insignificantly different from zero, but then increases to 0.108 i.e. in favour of native women and then further in favour of natives to 0.165 in 2011. These overall explained and unexplained components are precisely estimated, with standard errors of 0.1-0.4.

Investigating the second panel of Table 3.11, we can see what variables that most of the gap derives from for the explained part of the decomposition. There are only two variables that are consistently significantly different from zero over the three years, and they contribute in opposite directions to each other. The first of these initial endowments, and one that favours the migrant cohort, is for the amount of education that the two groups have, on average. The estimated contribution of the gap in the endowment of years of education in 2001 is equal to -0.166, falling slightly to -0.149 in 2006 and -0.142 in 2011. All three of these estimated gaps in the endowments of education are significant at the 0.1% level.

The other variable that is consistently significantly different from zero at the 0.1% level is the endowment of tenure. Here the contribution consistently favours natives, and to an increasing extent. This variable starts at 0.0153 in 2001, increasing to 0.0291 in 2006 and to 0.0405 in 2011.

I now consider the unexplained gaps. Again, the key variables with coefficients that are estimated precisely and are different from zero are education and tenure. The gap relating to the return to tenure is not significantly different from zero in 2001, but the gaps become significant and negative in 2006 and 2011 at -0.0322 and -0.0537, respectively. These gaps both favour the migrant population over the native population. The pattern with respect to education shows larger movements between periods – a very large contribution in favour of natives in 2001, then a contribution that is insignificantly different from zero in 2006, then another large contribution in 2011. This is attributable in part to imprecise estimates and large standard errors.

3.6.3.3 Threefold Decomposition

We can now move on to the results for the threefold decomposition in Table 3.12. We have, again, the details of the average hourly wages that are received by both the native and the migrant groups in the top panel of this table, now decomposed into contributions

of endowments, coefficients, and interactions. In 2001, the -0.121 log point's gap in favour of female migrants is largely attributable to the interaction of endowments and coefficients, at -0.110 log points; neither endowments nor coefficients on their own contribute anything significantly different from zero. The main difference in 2006 and 2011 is the increasing contribution of coefficients in favour of female migrants, to 0.108 in 2006 and 0.165 in 2011, both precisely determined and significantly different from zero. The contributions of endowments and endowment-coefficient interactions in favour of migrant women is enough to offset this in 2006, but by 2011 it the coefficient effect in favour of native women dominates both the endowment and interaction effects in favour of migrant women (the pure endowment effect is negligible in 2011).

If the gaps attributable to endowments in the second panel of Table 3.12 are examined, it can be seen that education endowments contribute in favour of migrant women: in all three years, foreign women have a higher level of education than native females, or have at least attended school for longer. Tenure, by contrast, contributes in favour of the native female population, and increasingly so: by 0.0203 log points in 2001, rising to 0.0941 in 2011.

The results for the gaps in the estimated coefficients are in the third panel of Table 3.12. We now need to review the results for the interaction between the gap in endowments and the gap in coefficients, presented in the bottom panel.

3.6.3.4 Counterfactuals

Counterfactuals were estimated, once again, for the female groupings over the study period. Once again a very similar story is told. The third row estimate the earnings that female migrants would earn if they were rewarded like their equivalent natives.

Table 3.13 shows that for the migrant female group, if they were rewarded for their endowments at the same rates as the native group they would be earning substantially more than they actually are earning. If the native female cohort had the same coefficients as the migrant group, they would be earning less than they currently do. For instance, in 2001 the estimated counterfactual for the migrant group would have been 1.96 log points (£7.10) instead of the 1.805 log points (£6.08) that they actually earned. In 2006 the female group of migrants were still be earning more than the native group of females, who earned 1.9 log points (£6.68). The counterfactual estimates that

they would have earned 2.08 log points or £8.00 if they had the same returns as the native group who only actually earned 1.881 log points (£6.56). This shows, like in all of the 2001 average wage level counterfactuals, that if the migrants were rewarded for their level of skill as the native cohort are, then the gap between the two pay levels would have grown.

3.7 Conclusion

This chapter has shown the reader that between the years of 2001 and 2011, the accession of the new member states to the European Union has dramatically changed the relative positions, at the mean, of the pay rates between native and migrant workers. Table 3.2 shows the reader that as the proportion of migrants in the UK being employed as Managers and Directors has fallen, from 11.6% in 2001 and by 2011 this number was only 4%. The opposite has occurred in Elementary occupations, with the proportion of the migrant sample employed in these jobs has increased from 16.9% going up to 23.1%. This change obviously took time to have an effect on the overall wages. As can be seen by 2006, in Tables 3.5 and 3.6, which looks at the overall average wages of natives and migrants, the migrant wage is surpassed by the native worker's average wage. However, the chapter then goes onto investigate whether this holds true for both genders.

The next few Tables then present the same analysis but independently for the two genders. Tables 3.8 and 3.9 show that the male populations of workers show much the same pattern as the overall decomposition. A different story emerges for the female populations as, even in 2006; the average native females are earning less than an equivalent migrant female. This shows that the two genders are exhibiting a difference to the male workers in terms of time taken, with native female wages taking a little bit longer to catch and then surpass the average wages of the female migrant population.

It would appear from the results of the decompositions that the most important aspects are an individual's education and the amount of tenure that they have accrued working in the job in which they are currently employed. The amount of education that an individual has is as important for the native population and for the migrant cohort before the expansion of the EU that began in 2004. We can see this as before the expansion of the EU, the levels of a migrants education easily outstrips the level of education that a native worker has.

This, however, does not continue after 2004 with the expansion of the EU brought a new wave of migration deriving, in most part, from the new EU member states. If Table 1a is examined it can be seen that the migrant population seems to continue to have 2 more years in education over what the average British person has. However the largest gap in education, between native and migrant cohorts, occurs in the male population in 2011 with a gap of 3.59 years in the average level of education, with migrant workers averaging 15.29 years and the average for the native population at only 12.7 years.

The information that is presented in this chapter shows how the expansion of the EU has had an effect on the relative position of native and migrants' workers. This can also be seen in the various subpopulations presented in Figure 2.1. This figure agrees with the analysis that I have presented in this chapter. It shows that the native male population was earning more than the equivalent migrant was in 2006 and that this gap persisted with an increase in the gap in the years after this analysis finished, after 2011. It does however have slightly different conclusions about the female population. The figures presented in Tables 3.11 and 3.12 show that the native female population is earning more than the migrant population by 2011.

Chapter Four

Conclusion

The second chapter applies the model developed by Borjas and Bratsberg, (1996), to data from the UK, in the period from 29/04/01 until 27/04/11. The analysis required the construction of a new and unique data set. This required combining a diverse range of other data sources including information from the 2011 UK census and information derived from the UK government's Department for Work and Pensions. These two data sources were used in order to calculate the rate of return migration for the 93 countries that were studied. The construction these data also allows the potential to look at the average selection that decides the type of migrant that comes from particular countries, on average.

I find that the two biggest factors that can affect outmigration rates, for different groups, are the distance between the source capital and London and the gap between the source and host countries GDP in 2011. The marginal effects of an increase in distance was found to be 692 more individuals returning home if the source country was 1000 miles further away from the UK. There was a larger reaction to the narrowing of the gap in the two countries GDP. It is found that if the gap closed by 1% there would be an additional 1091 individuals returning to the source country.

I also estimated the marginal effects of the change in variables for the two largest migrant cohorts in the UK over the time being selected for study, Poland and India. These two countries reinforce many of the points from the initial paper by Borjas and Bratsberg. Distance from London, which was selected as a proxy for travel costs, to New Dehli is over 5000 miles more than it is to Warsaw, and is therefore considered more expensive to travel from/to. Almost twice as many people are going to travel between the capital of India (10904) and London than would between the London and Warsaw (5760).

Once we look at the two countries differing reactions to a decrease in the gap between the GDP of host and source country it can is observed that the same reaction is being experienced by migrants of India and Poland. If there is a decrease in the GDP gap between India and the UK this is estimated to induce 17204 more individuals to return to India. The same situation happening between the UK and Poland and is estimated to lead to 9087 more Poles returning to their source country.

In the Third chapter, I show the reader that between the years of 2001 and 2011, the accession of the new states to the European Union has dramatically changed the relative

positions, at the mean, of the pay rates between native and migrant workers. Table 3.2 shows the reader that as the proportion of migrants in the UK being employed as Managers and Directors has fallen, from 11.6% in 2001 by 2011 this number was only 4%. The opposite has occurred in Elementary occupations, with the proportion of the migrant sample employed in these jobs has increased from 16.9% up to 23.1%. This change obviously took time to have an effect on the overall wages. As can be seen by 2006, in Tables 5 and 6, which looks at the overall average wages of natives and migrants, the migrant wage is surpassed by the native worker's average wage. However, the chapter then goes onto investigate whether this holds true for both genders.

The next few Tables then present the same analysis but independently for the two genders. Tables 3.8 and 3.9 show that the male populations of workers show much the same pattern as the overall decomposition. A different story emerges for the female populations as, even in 2006; native females are earning less than a migrant female. This shows that the two genders are exhibiting a major difference to the male workers, with native female wages taking a little bit longer to catch and then surpass the average wages of the female migrant population.

It would appear from the results of the decompositions that the most important aspects are an individual's education and the amount of tenure that they have accrued working in the job in which they are currently employed. The amount of education that an individual has can be seen as important for the native population and for the migrant cohort before the expansion of the EU that began in 2004. We can see this as before the expansion of the EU, the levels of a migrants education easily outstrips the level of education that a native worker has.

This, however, does not continue after 2004 with the expansion of the EU brought a new wave of migration deriving in most part from the new EU member states. If Table 1a is examined it can be seen that the migrant population seems to continue to have 2 more years in education over what the average British person has. However the largest gap in education, between native and migrant cohorts, occurs in the male population in 2011 with a gap of 3.59 years in the average level of education, with migrant workers averaging 15.29 years and the average for the native population at only 12.7 years.

The information that is presented in this chapter shows how the expansion of the EU has had an effect on the relative position of native and migrants' workers. This can

also be seen in the various subpopulations presented in Figure 2.1. This figure agrees with the analysis that I have presented in this chapter. It shows that the native male population were earning more than the equivalent migrant was in 2006 and that this gap persisted with a small increase in the gap in the years after this analysis finished, after 2011. It does however have slightly different conclusions about the female population. The figures presented in Tables 3.11 and 3.12 show that the native female population is earning more than the migrant population by 2011.

Additional avenues for further research include investigating wage decomposition based on individuals' occupation the same way that Brown, Moon and Zoloth (1980) did in their original paper. This would allow for the analysis of how occupational differences between native and migrant have an effect on the wages earned by the two groups.

Appendix A:

Graphs and Tables – Chapter Two

Table A1 – Determinants of Outmigration and In-Migration Rates

Variable	Outmigration Rate				In-Migration Rate			
	GLS			Grouped Probit	GLS			Grouped Probit
	(1)	(2)	(3)		(1)	(2)	(3)	
Intercept	0.3900 (13.18)	0.2371 (4.30)	0.3283 (5.43)	-0.4727 (-2.39)	3.4580 (6.47)	2.2680 (2.42)	4.5401 (2.08)	-2.7700 (-38.16)
Log (per capita GNP)	0.0249 (2.07)	0.0488 (3.82)	0.0443 (3.07)	0.2960 (6.22)	-0.3004 (-2.04)	-0.2268 (-1.44)	-0.4904 (-0.94)	-0.0173 (-0.76)
Distance (in 000's miles)	-0.0277 (-2.02)	-0.0140 (-2.03)	-0.0140 (-2.02)	-0.1136 (-5.45)	-0.5678 (-5.10)	-0.5412 (-4.33)	-1.0723 (-4.28)	-0.117 (-13.84)
Communist Regime	-0.2122 (-6.10)	-0.1148 (-2.66)	-0.1820 (-3.60)	-0.9841 (-6.06)	-0.7519 (-2.53)	-0.2806 (-0.68)	-0.4731 (-0.26)	-0.1687 (-3.99)
Coup or Revolution	-0.0089 (-0.24)	-0.0183 (-0.50)	-0.0564 (-1.48)	-0.0721 (-0.58)	0.0505 (0.09)	-0.1813 (-0.30)	-2.0419 (-1.49)	-0.0095 (-0.19)
Income Inequality	…	0.0177 (2.52)	0.0084 (0.93)	0.0649 (2.28)	…	0.3053 (2.07)	0.5527 (1.69)	0.0272 (2.40)
Income Inequality Squared	…	-0.0003 (-1.37)	-0.0001 (-0.36)	-0.0013 (-1.41)	…	-0.0109 (-2.22)	-0.067 (-1.69)	-0.0007 (-2.03)
1970-1974 cohort	-0.0349 (-1.50)	-0.0257 (-1.17)	-0.0297 (-1.05)	0.0443 (0.58)	-0.0333 (-0.12)	-0.0693 (-0.25)	0.2205 (0.22)	0.0290 (0.85)
R^2	0.439	0.512	0.310	…	0.242	0.276	0.242	…
Weighted	Yes	Yes	No	Yes	Yes	Yes	No	Yes

Note: the t-ratios are reported in parenthesis. The weighted regressions weigh the observations by the size of the immigrant flow in the outmigration equation, and by the size of the source countries population in the in-migration regression. The regressions have 119 observations.

Table 2.1 – Regression upon in and out migration rates

	Out Migration			Immigration		
	Grouped Probit	Grouped Probit With robust standard errors	WLS	Grouped Probit	Grouped Probit With robust standard errors	WLS
Distance in 000s miles	0.0495***	0.0495**	0.0490***	-0.0772***	-0.0772	-0.112***
	(0.000292)	(0.0161)	(0.0122)	(0.0000951)	(0.0477)	(0.0171)
GDP Gap in 2011	0.0781***	0.0781	0.0672	-0.164***	-0.164	-0.205**
	(0.00182)	(0.0813)	(0.0763)	(0.000376)	(0.0915)	(0.0755)
Political Stability	0.00333***	0.00333	0.00312	-0.0157***	-0.0157**	-0.0147***
	(0.0000835)	(0.00309)	(0.00350)	(0.0000199)	(0.00597)	(0.00381)
Income Inequality	-0.00832***	-0.00832	-0.00840	-0.00130***	-0.00130	-0.000329
	(0.000239)	(0.0107)	(0.0101)	(0.0000564)	(0.0163)	(0.0104)
Income Inequality^2	0.000128***	0.000128	0.000131	0.00000701***	0.00000701	0.0000443
	(0.00000352)	(0.000141)	(0.000149)	(0.000000837)	(0.000211)	(0.000153)
Constant	-0.312***	-0.312*	-0.314*	-2.215***	-2.215***	-2.407***
	(0.00323)	(0.153)	(0.136)	(0.000728)	(0.260)	(0.134)
N	3244601	3244601	91	5958273174	5958273174	91

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.2 – Marginal Effects of independent variables

Variables	Maximum Likelihood			Maximum Likelihood with Cluster Robust Estimation of the Variance			Weighted Least Squares		
	Marginal Effects Overall (at the Mean)	Marginal Effects for Poles (at the Mean)	Marginal Effects for Indians (at the Mean)	Marginal Effects Overall (at the Mean)	Marginal Effects for Poles (at the Mean)	Marginal Effects for Indians (at the Mean)	Marginal Effects Overall (at the Mean)	Marginal Effects for Poles (at the Mean)	Marginal Effects for Indians (at the Mean)
Distance (000's of Miles)	691.57 (4.02)	10903.83 (63.29)	5759.51 (34.13)	691.57 (221.84)	10903.83 (3499.06)	5759.51 (1878.88)	684.57 (167.66)	10785.75 (2636.57)	5707.12 (1424.14)
GDP Gap 2011	1091.12 (25.37)	17203.53 (399.75)	9087.08 (210.99)	1091.12 (1136.15)	17203.53 (17921.78)	9087.08 (9422.06)	937.82 (1064.53)	144775.83 (16758.8)	7818.41 (8862.92)
Political Stability	46.57 (1.17)	734.34 (18.34)	387.89 (9.7)	46.57 (43.15)	734.34 (679.39)	387.89 (358.13)	43.48 (1064.53)	685.08 (768.06)	362.5 (406.9)
Income Inequality	-73.9 (2.28)	-1336.13 (40.13)	-750.55 (22.33)	-116.18 (149.86)	-1831.87 (2365.35)	-967.61 (1252.58)	-73.98 (48.82)	-1340.64 (1686.17)	-755.0 (939.92)

Standard errors in parentheses

Figure 2.1 (Positive Selection)

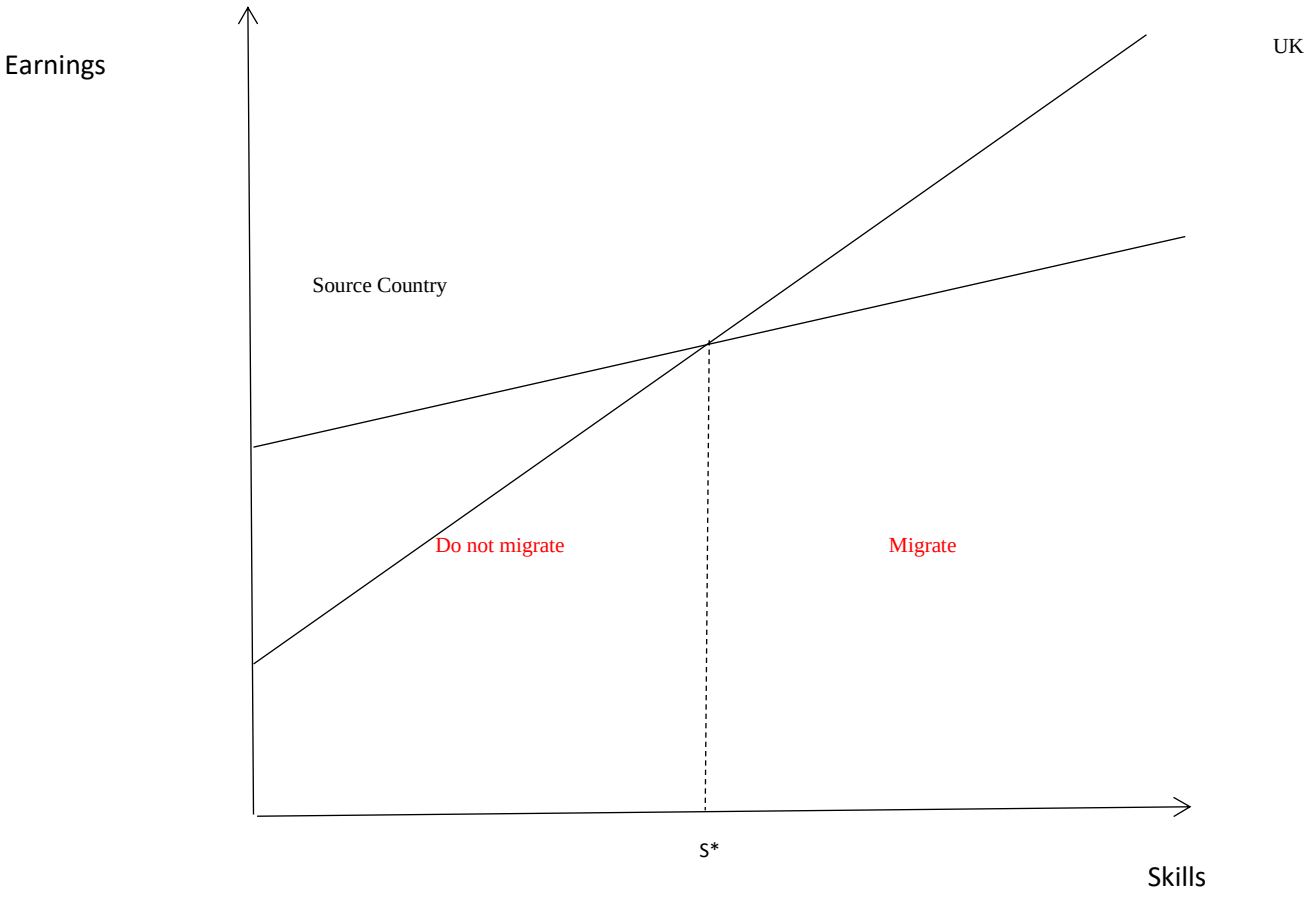


Figure 2.2 (Negative Selection)

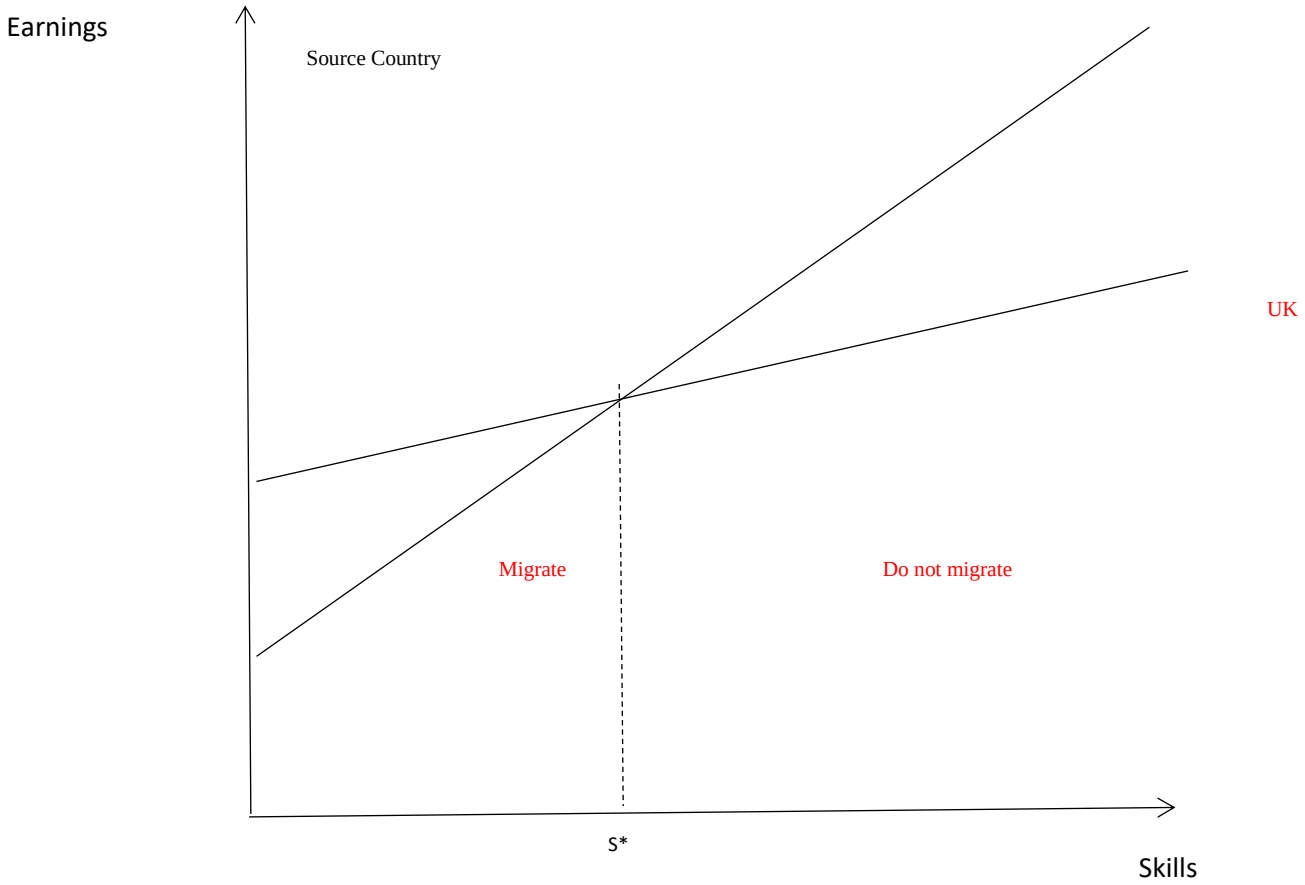
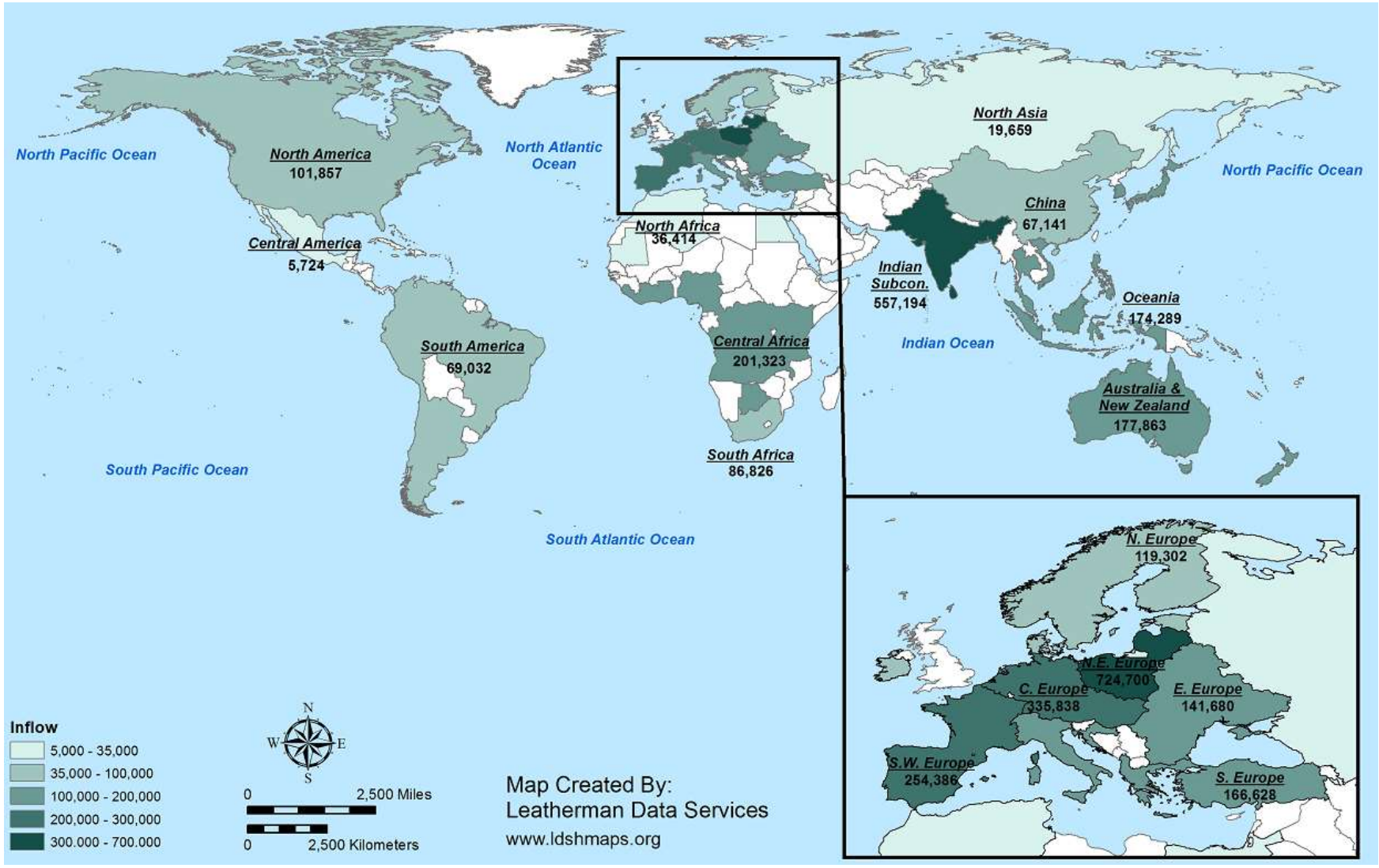


Figure 2.3 World Population Inflow to the UK



Appendix B:

Graphs and Tables – Chapter Three

Table 3.1 – Descriptive Statistics for Natives and Migrants (totals and by gender)

a) 3.1.1 Total Sample of Natives & Migrants

Men	2001		2006		2011	
	Native	Migrant	Native	Migrant	Native	Migrant
Hourly wage, mean	6.56 (4.44)	7.76 (7.66)	8.14 (6.78)	8.3 (5.9)	9.3 (5.8)	8.9 (6.1)
Hourly wage, median	5.56	6.00	6.81	6.67	7.84	7.76
Log hourly wage, mean	1.76 (0.49)	1.84 (0.59)	1.97 (0.48)	1.95 (0.56)	2.1 (0.5)	2 (0.5)
Log hourly wage, median	1.71	1.79	1.97	1.95	2.1	1.79
Age	40.95 (11.5)	39.08 (10.8)	43.8 (10.2)	38.58 (9.65)	44.4 (10.3)	38.73 (9.57)
Education	12.1 (2.39)	14.23 (3.76)	12.28 (2.51)	14.27 (3.61)	12.58 (2.62)	14.78 (3.43)
Experience	22.11 (12.39)	18.32 (12.39)	20.52 (10.95)	13.3 (10.76)	20.87 (11.09)	12.95 (10.49)
Tenure	7.39 (8.24)	4.7 (6.32)	8.7 (8.62)	4.2 (5.58)	9.52 (8.8)	4.57 (5.15)
Married	0.59 (0.49)	0.62 (0.48)	0.34 (0.18)	0.62 (0.24)	0.35 (0.18)	0.47 (0.21)
N=	10782	434	12655	731	11498	1033

Standard errors in parentheses

b) 3.1.2 Male Natives & Migrants

Men	2001		2006		2011	
	Native	Migrant	Native	Migrant	Native	Migrant
Hourly wage, mean	7.35 (0.08)	8.19 (0.53)	9.09 (0.13)	8.63 (0.3)	10.08 (0.09)	9.48 (0.28)
Hourly wage, median	1.77	1.63	6.5	6.24	7.73	7.1
Log hourly wage, mean	1.86 (0.008)	1.87 (0.046)	2.07 (0.01)	1.99 (0.03)	2.17 (0.00)	2.1 (0.2)
Log hourly wage, median	5.85	5.11	1.87	1.83	2.05	1.96
Age	41 (0.2)	37.9 (0.77)	44 (0.15)	37.1 (0.47)	44.7 (0.16)	37.9 (0.42)
Education	12.18 (0.04)	14.03 (0.27)	12.5 (0.4)	14.8 (0.2)	12.7 (0.04)	15.29 (0.16)
Experience	15.3 (0.19)	13.63 (0.71)	17.21 (0.17)	12.9 (0.49)	16.9 (0.17)	12.5 (0.43)
Tenure	7.56 (0.14)	4.25 (0.43)	8.4 (0.12)	3.43 (0.26)	9.02 (0.12)	4.1 (0.21)
Married	0.61 (0.01)	0.64 (0.03	0.66 (0.00)	0.625 (0.02)	0.63 (0.00)	0.61 (0.02)
N=	5992	217	7149	397	6299	592

Standard errors in parentheses

c) 3.1.3 Female Natives & Migrants

Women	2001		2006		2011	
	Native	Migrant	Native	Migrant	Native	Migrant
Hourly wage, mean	5.97 (0.4)	7.26 (0.54)	7.28 (0.05)	7.7 (0.27)	8.54 (0.06)	8.28 (0.23)
Hourly wage, median	6.25	6.16	6.73	6.54	8.51	7.51
Log hourly wage, mean	1.68 (0.01)	1.81 (0.04)	1.88 (0.01)	1.90 (0.27)	2.04 (0.01)	1.98 (0.02)
Log hourly wage, median	1.83	1.82	1.91	1.88	2.14	2.02
Age	40.89 (0.15)	40.12 (0.74)	43.52 (0.12)	38.35 (0.19)	44.24 (0.13)	37.7 (0.40)
Education	11.89 (0.03)	14.26 (0.26)	12.18 (0.03)	14.53 (0.19)	12.49 (0.03)	14.94 (0.14)
Experience	16.61 (0.14)	14.99 (0.43)	17.88 (0.13)	13.79 (0.49)	17.36 (0.14)	12.57 (0.4)
Tenure	6.39 (0.09)	4.87 (0.43)	7.47 (0.09)	4.03 (0.28)	8.39 (0.1)	4.2 (0.2)
Married	0.6 (0.00)	0.62 (0.03)	0.61 (0.01)	0.62 (0.02)	0.58 (0.01)	0.6 (0.02)
N=	3920	197	5016	384	4719	526

Standard errors in parentheses

Table 3.2 - Occupational Distribution of the Different Groups

	2001		2006		2011	
	2001		2006		2011	
Occupation	Natives	Migrants	Natives	Migrants	Natives	Migrants
Managers & Directors	1074 (10.8%)	48 (11.6%)	1546 (12.7%)	83 (10.6%)	904 (8.2%)	45 (4%)
Professional Occupations	798 (8.1%)	63 (8.1%)	1259 (10.4%)	120 (15.4%)	1829 (16.6)	219 (19.6%)
Associate Professionals	1319 (13.3%)	49 (11.8%)	1682 (13.8%)	110 (14.1%)	1460 (13.3%)	105 (9.4%)
Administrative & Secretarial	1930 (19.47)	63 (15.2%)	2092 (17.2%)	68 (8.7%)	1796 (16.3%)	114 (10.2%)
Skilled Trades	711 (7.2%)	28 (6.8%)	838 (6.9%)	48 (6.2%)	800 (7.3%)	79 (7%)
Personal Service	953 (9.6%)	39 (9.42%)	1274 (10.5%)	83 (10.6%)	1217 (11.1%)	116 (10.4%)
Sales & Customer	874 (8.8%)	26 (6.3%)	979 (8%)	39 (5%)	987 (9%)	69 (6.2%)
Process & Plant	804 (8.1%)	28 (6.8%)	907 (7.5%)	65 (8.3%)	680 (6.2%)	113 (10.1%)
Elementary	1449 (14.6)	70 (16.9%)	1588 (13.1%)	156 (21.1%)	1345 (12.2%)	258 (23.1%)
N=	9912	414	12165	781	11018	1118

This table shows the raw numbers employed in the different occupations. The percentages that are reported in parenthesis are the percentages for each individual group.

Table 3.3 – Percentage that have Migrated from various World Regions

	2001	2006	2011
Western EU	23.6	27	21.1
Eastern EU	1.1	9.8	29.3
Europe non-EU	5	1.8	2.2
North America	7.6	4.9	3.5
South America	2.3	1.1	1.9
Australasia	9	5.3	2.1
Africa	18.8	15.2	12.5
Asia non Middle East	18.6	19.3	19.9
Middle East	2	1.8	1.4
Caribbean	4.2	2.8	1.7

Numbers may not sum to 100 due to rounding.

Source: UK Labour Force Survey

Table 3.4 – OLS results for Natives and Migrants (Total Populations)

	2001		2006		2011	
	Natives	Migrants	Natives	Migrants	Natives	Migrants
Education	0.0718***	0.0491***	0.0617***	0.0554***	0.0602***	0.0433***
	(0.00203)	(0.00843)	(0.00172)	(0.00563)	(0.00185)	(0.00472)
Experience	0.0146***	0.00244	0.00481***	0.00254	0.00705***	0.00619
	(0.00130)	(0.00941)	(0.00127)	(0.00639)	(0.00137)	(0.00492)
Experience^2	-0.000360***	-0.0000427	-0.000160***	-0.0000544	-0.000185***	-0.000151
	(0.0000325)	(0.000239)	(0.0000297)	(0.000165)	(0.0000330)	(0.000132)
Tenure	0.0115***	0.0133**	0.00806***	0.0231***	0.00975***	0.0229***
	(0.000612)	(0.00459)	(0.000564)	(0.00360)	(0.000604)	(0.00299)
Married	0.0458***	-0.00360	0.0715***	0.0259	0.0697***	0.0426
	(0.00964)	(0.0601)	(0.00840)	(0.0372)	(0.00913)	(0.0294)
Female	-0.150***	-0.0839	-0.162***	-0.0903*	-0.116***	-0.106***
	(0.00899)	(0.0560)	(0.00803)	(0.0357)	(0.00878)	(0.0287)
constant	0.779***	1.105***	1.171***	1.056***	1.235***	1.280***
	(0.0304)	(0.159)	(0.0296)	(0.116)	(0.0320)	(0.0950)
Marg. eff.	0.0027	-0.0001	-0.0002	0.0017	0.0007	0.0018
of experience	(0.0005)	(0.0036)	(0.0004)	(0.0022)	(0.0005)	(0.0021)
N	9912	414	12165	781	11018	1118

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.5 – Twofold Wage Decomposition with Natives Non-Discriminatory Wage Regression

	2001	2006	2011
Overall			
Native	1.752***	1.958***	2.097***
	(0.00481)	(0.00432)	(0.00468)
Migrant	1.835***	1.944***	2.037***
	(0.0292)	(0.0195)	(0.0153)
Difference	-0.0826**	0.0144	0.0593***
	(0.0296)	(0.0200)	(0.0160)
Explained	-0.142***	-0.130***	-0.111***
	(0.0145)	(0.0104)	(0.00870)
unexplained	0.0598*	0.144***	0.170***
	(0.0291)	(0.0193)	(0.0163)
Explained			
Education	-0.154***	-0.145***	-0.152***
	(0.0143)	(0.0101)	(0.00874)
Experience	-0.00139	-0.00557**	0.00193
	(0.00246)	(0.00202)	(0.00240)
Tenure	0.0261***	0.0332***	0.0439***
	(0.00391)	(0.00284)	(0.00308)
Married	-0.00103	0.000762	-0.000366
	(0.00113)	(0.00128)	(0.00107)
Female	-0.0121**	-0.0129***	-0.00489**
	(0.00384)	(0.00306)	(0.00186)
Coefficients			
Education	0.322*	0.0923	0.256**
	(0.143)	(0.0880)	(0.0885)
Experience	0.0760	0.00152	0.00221
	(0.0747)	(0.0460)	(0.0331)
Tenure	-0.00822	-0.0560***	-0.0547***
	(0.0205)	(0.0123)	(0.0143)
Married	0.0311	0.0283	0.0164
	(0.0374)	(0.0233)	(0.0178)
Female	-0.0349	-0.0365	-0.00511
	(0.0304)	(0.0188)	(0.0158)
_cons	-0.326	0.115	-0.0444
	(0.190)	(0.118)	(0.111)
N	10326	12946	12136

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.6 – Threefold Wage Decomposition

	2001	2006	2011
Overall			
Natives	1.752***	1.958***	2.097***
	(0.00481)	(0.00432)	(0.00468)
Migrants	1.835***	1.944***	2.037***
	(0.0292)	(0.0195)	(0.0153)
Difference	-0.0826**	0.0144	0.0593***
	(0.0296)	(0.0200)	(0.0160)
endowments	-0.0806***	-0.0406*	-0.00703
	(0.0224)	(0.0178)	(0.0192)
coefficients	0.0598*	0.144***	0.170***
	(0.0291)	(0.0193)	(0.0163)
interaction	-0.0617**	-0.0893***	-0.104***
	(0.0216)	(0.0169)	(0.0192)
Explained			
Endowments	-0.0806***	-0.0406*	-0.00703
	(0.0224)	(0.0178)	(0.0192)
Interaction	-0.0617**	-0.0893***	-0.104***
	(0.0216)	(0.0169)	(0.0192)
Endowments			
Education	-0.105***	-0.131***	-0.109***
	(0.0230)	(0.0152)	(0.0144)
Experience	0.00107	0.00193	0.00356
	(0.00499)	(0.00781)	(0.00773)
Tenure	0.0302**	0.0950***	0.103***
	(0.0109)	(0.0138)	(0.0156)
Married	0.0000809	0.000276	-0.000224
	(0.00132)	(0.000605)	(0.000670)
Female	-0.00674	-0.00716*	-0.00449*
	(0.00505)	(0.00330)	(0.00206)
Coefficients			
Education	0.322*	0.0923	0.256**
	(0.143)	(0.0880)	(0.0885)
Experience	0.0760	0.00152	0.00221
	(0.0747)	(0.0460)	(0.0331)
Tenure	-0.00822	-0.0560***	-0.0547***
	(0.0205)	(0.0123)	(0.0143)
Married	0.0311	0.0283	0.0164
	(0.0374)	(0.0233)	(0.0178)
Female	-0.0349	-0.0365	-0.00511
	(0.0304)	(0.0188)	(0.0158)
constant	-0.326	0.115	-0.0444
	(0.190)	(0.118)	(0.111)
Interaction			
Education	-0.0487*	-0.0148	-0.0427**
	(0.0221)	(0.0142)	(0.0149)
Experience	-0.00246	-0.00750	-0.00163
	(0.00540)	(0.00806)	(0.00805)
Tenure	-0.00408	-0.0618***	-0.0593***
	(0.0102)	(0.0135)	(0.0155)
Married	-0.00111	0.000486	-0.000143
	(0.00179)	(0.000909)	(0.000444)
Female	-0.00535	-0.00571	-0.000407
	(0.00494)	(0.00321)	(0.00127)
N	10326	12946	12136

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.7 – OLS Results for Male Sample of Natives & Migrants

	2001		2006		2011	
	Natives	Migrants	Natives	Migrants	Natives	Migrants
Education	0.0721***	0.0729***	0.0596***	0.0573***	0.0623***	0.0496***
	(0.00309)	(0.0123)	(0.00261)	(0.00813)	(0.00277)	(0.00645)
Experience	0.0243***	-0.00566	0.0136***	-0.00543	0.0134***	0.0167*
	(0.00215)	(0.0147)	(0.00207)	(0.00892)	(0.00220)	(0.00727)
Experience^2	-0.000575***	0.000227	-0.000375***	0.000125	-0.000306***	-0.000396*
	(0.0000522)	(0.000368)	(0.0000474)	(0.000221)	(0.0000516)	(0.000191)
Tenure	0.0129***	0.0115	0.00791***	0.0326***	0.0101***	0.0236***
	(0.000985)	(0.00715)	(0.000885)	(0.00585)	(0.000945)	(0.00450)
Married	0.123***	-0.0363	0.144***	0.0445	0.0938***	0.000896
	(0.0170)	(0.0930)	(0.0143)	(0.0538)	(0.0151)	(0.0442)
constant	0.650***	0.832***	1.093***	1.040***	1.134***	1.137***
	(0.0460)	(0.228)	(0.0451)	(0.166)	(0.0480)	(0.131)
Marg. eff.	0.0721	0.0729	0.0007	-0.0022	0.003	0.007
of experience	(0.0035)	(0.0137)	(0.0007)	(0.0041)	(0.0008)	(0.0023)
N	3920	197	5016	384	4719	526

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.8 – Twofold Wage Decomposition Male Natives & Migrants with Natives as Non-discriminatory Group

	2001	2006	2011
overall			
Natives	1.857***	2.068***	2.174***
	(0.00833)	(0.00706)	(0.00754)
Migrants	1.868***	1.989***	2.101***
	(0.0464)	(0.0284)	(0.0229)
difference	-0.0106	0.0792**	0.0735**
	(0.0472)	(0.0293)	(0.0241)
explained	-0.104***	-0.100***	-0.101***
	(0.0218)	(0.0143)	(0.0133)
unexplained	0.0931*	0.180***	0.175***
	(0.0445)	(0.0279)	(0.0246)
Explained			
Education	-0.133***	-0.138***	-0.161***
	(0.0204)	(0.0141)	(0.0137)
Experience	-0.0105	-0.00711	0.00827*
	(0.00562)	(0.00388)	(0.00395)
Tenure	0.0428***	0.0393***	0.0498***
	(0.00680)	(0.00486)	(0.00488)
Married	-0.00307	0.00498	0.00154
	(0.00433)	(0.00372)	(0.00211)
Coefficients			
Education	-0.0113	0.0329	0.194
	(0.199)	(0.120)	(0.133)
Experience	0.179	0.116	-0.0183
	(0.119)	(0.0668)	(0.0531)
Tenure	0.00608	-0.0845***	-0.0552**
	(0.0324)	(0.0188)	(0.0208)
Married	0.102	0.0620	0.0571*
	(0.0630)	(0.0340)	(0.0280)
constant	-0.182	0.0530	-0.00322
	(0.277)	(0.161)	(0.165)
N	4117	5400	5245

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.9 - Threefold Oaxaca- Blinder Decomposition for Male Natives & Migrants

	2001	2006	2011
Overall			
Natives	1.857***	2.068***	2.174***
	(0.00833)	(0.00706)	(0.00754)
Migrants	1.868***	1.989***	2.101***
	(0.0464)	(0.0284)	(0.0229)
difference	-0.0106	0.0792**	0.0735**
	(0.0472)	(0.0293)	(0.0241)
endowments	-0.0848*	0.0295	-0.00436
	(0.0383)	(0.0319)	(0.0291)
coefficients	0.0931*	0.180***	0.175***
	(0.0445)	(0.0279)	(0.0246)
interaction	-0.0189	-0.130***	-0.0968***
	(0.0350)	(0.0306)	(0.0292)
Explained			
Endowment	-0.0848*	0.0295	-0.00436
	(0.0383)	(0.0319)	(0.0291)
Interaction	-0.0189	-0.130***	-0.0968***
	(0.0350)	(0.0306)	(0.0292)
Endowments			
Education	-0.134***	-0.133***	-0.128***
	(0.0320)	(0.0209)	(0.0224)
Experience	0.0107	-0.00149	0.00778
	(0.00943)	(0.0121)	(0.0113)
Tenure	0.0380	0.162***	0.116***
	(0.0256)	(0.0269)	(0.0250)
Married	0.000908	0.00155	0.0000147
	(0.00273)	(0.00214)	(0.000703)
Coefficients			
Education	-0.0113	0.0329	0.194
	(0.199)	(0.120)	(0.133)
Experience	0.179	0.116	-0.0183
	(0.119)	(0.0668)	(0.0531)
Tenure	0.00608	-0.0845***	-0.0552**
	(0.0324)	(0.0188)	(0.0208)
Married	0.102	0.0620	0.0571*
	(0.0630)	(0.0340)	(0.0280)
constant	-0.182	0.0530	-0.00322
	(0.277)	(0.161)	(0.165)
Interaction			
Education	0.00149	-0.00515	-0.0328
	(0.0261)	(0.0188)	(0.0226)
Experience	-0.0212	-0.00562	0.000494
	(0.0115)	(0.0129)	(0.0115)
Tenure	0.00475	-0.123***	-0.0660**
	(0.0253)	(0.0267)	(0.0249)
Married	-0.00398	0.00344	0.00152
	(0.00610)	(0.00316)	(0.00221)
N	4117	5400	5245

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.10 – OLS results for Female Natives & Migrants

	2001		2011		2001	
	Natives	Migrants	Natives	Migrants	Natives	Migrants
Education	0.0701***	0.0232*	0.0631***	0.0530***	0.0578***	0.0356***
	(0.00268)	(0.0115)	(0.00230)	(0.00787)	(0.00249)	(0.00696)
Experience	0.00793***	0.00624	-0.00183	0.0130	0.00237	-0.00379
	(0.00163)	(0.0120)	(0.00160)	(0.00934)	(0.00176)	(0.00673)
Experience^2	-0.000214***	-0.000263	0.000000683	-0.000313	-0.0000987*	0.0000882
	(0.0000412)	(0.000309)	(0.0000379)	(0.000249)	(0.0000429)	(0.000183)
Tenure	0.0102***	0.0135*	0.00847***	0.0165***	0.00965***	0.0224***
	(0.000782)	(0.00585)	(0.000732)	(0.00458)	(0.000782)	(0.00399)
Married	-0.00188	0.0141	0.0251*	0.00403	0.0509***	0.0771
	(0.0114)	(0.0779)	(0.0102)	(0.0520)	(0.0113)	(0.0395)
constant	0.740***	1.395***	1.066***	0.971***	1.206***	1.334***
	(0.0390)	(0.221)	(0.0380)	(0.160)	(0.0418)	(0.134)
Marg. eff.	0.0008	-0.0017	-0.0018	0.0044	-0.0011	-0.0016
of experience	(0.0006)	(0.005)	(0.00058)	(0.0034)	(0.0006)	(0.0026)
N	5992	217	7149	397	6299	592

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.11 - Twofold Oaxaca- Blinder Decomposition for Female Natives & Migrants

	2001	2006	2011
overall			
Natives	1.684***	1.881***	2.038***
	(0.00563)	(0.00524)	(0.00583)
Migrants	1.805***	1.900***	1.981***
	(0.0369)	(0.0268)	(0.0204)
difference	-0.121**	-0.0191	0.0574**
	(0.0374)	(0.0273)	(0.0212)
explained	-0.151***	-0.127***	-0.107***
	(0.0192)	(0.0140)	(0.0108)
unexplained	0.0295	0.108***	0.165***
	(0.0389)	(0.0270)	(0.0217)
Explained			
Education	-0.166***	-0.149***	-0.142***
	(0.0198)	(0.0143)	(0.0110)
Experience	0.0000967	-0.00737**	-0.00503
	(0.00204)	(0.00250)	(0.00311)
Tenure	0.0153***	0.0291***	0.0405***
	(0.00465)	(0.00359)	(0.00401)
Married	0.0000347	-0.000129	-0.00103
	(0.000219)	(0.000631)	(0.00110)
Coefficients			
Education	0.669***	0.147	0.331**
	(0.198)	(0.128)	(0.117)
Experience	0.0417	-0.115	0.0297
	(0.0915)	(0.0648)	(0.0415)
Tenure	-0.0162	-0.0322*	-0.0537**
	(0.0270)	(0.0164)	(0.0197)
Married	-0.00992	0.0130	-0.0157
	(0.0431)	(0.0319)	(0.0229)
constant	-0.655*	0.0952	-0.127
	(0.259)	(0.174)	(0.146)
N	6209	7546	6891

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.12 - Threefold Oaxaca- Blinder Decomposition for Female Natives & Migrants

	2001	2006	2011
Overall			
Natives	1.684***	1.881***	2.038***
	(0.00563)	(0.00524)	(0.00583)
Migrants	1.805***	1.900***	1.981***
	(0.0369)	(0.0268)	(0.0204)
difference	-0.121**	-0.0191	0.0574**
	(0.0374)	(0.0273)	(0.0212)
endowments	-0.0404	-0.0621**	0.00176
	(0.0296)	(0.0209)	(0.0250)
coefficients	0.0295	0.108***	0.165***
	(0.0389)	(0.0270)	(0.0217)
interaction	-0.110***	-0.0648**	-0.109***
	(0.0314)	(0.0203)	(0.0252)
Explained			
Endowments	-0.404	-0.0621**	0.00176
	(0.0296)	(0.0209)	(0.0250)
Interaction	-0.110	-0.0648**	-0.109***
	(0.0314)	(0.0203)	(0.0252)
Endowments			
Education	-0.0549	-0.125***	-0.0872***
	(0.0324)	(0.0217)	(0.0183)
Experience	-0.00559	0.00608	-0.00353
	(0.00677)	(0.00999)	(0.0103)
Tenure	0.0203*	0.0566***	0.0941***
	(0.0101)	(0.0145)	(0.0198)
Married	-0.000260	-0.0000208	-0.00156
	(0.00135)	(0.000280)	(0.00179)
Coefficients			
Education	0.669***	0.147	0.331**
	(0.198)	(0.128)	(0.117)
Experience	0.0417	-0.115	0.0297
	(0.0915)	(0.0648)	(0.0415)
Tenure	-0.0162	-0.0322*	-0.0537**
	(0.0270)	(0.0164)	(0.0197)
Married	-0.00992	0.0130	-0.0157
	(0.0431)	(0.0319)	(0.0229)
constant	-0.655*	0.0952	-0.127
	(0.259)	(0.174)	(0.146)
Interaction			
Education	-0.111**	-0.0238	-0.0543**
	(0.0350)	(0.0208)	(0.0194)
Experience	0.00568	-0.0134	-0.00150
	(0.00655)	(0.0104)	(0.0107)
Tenure	-0.00500	-0.0275	-0.0536**
	(0.00847)	(0.0141)	(0.0197)
Married	0.000295	-0.000108	0.000531
	(0.00139)	(0.000591)	(0.000954)
N	6209	7546	6891

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.13 – Predicted Average Wage Counterfactuals

	Type of Counterfactual	2001	2006	2011
Overall Groups	Migrant Endowments, Native Coefficients	1.88 (0.28)	2.07 (0.26)	2.19 (0.23)
	Native Endowments, Migrant Coefficients	1.77 (0.19)	1.93 (0.27)	2.04 (0.24)
Male Groups	Migrant Endowments, Native Coefficients	1.83 (0.26)	2.01 (0.24)	2.15 (0.20)
	Native Endowments, Migrant Coefficients	1.77 (0.13)	1.84 (0.16)	1.98 (0.20)
Female Groups	Migrant Endowments, Native Coefficients	1.96 (0.28)	2.16 (0.24)	2.28 (0.2)
	Native Endowments, Migrant Coefficients	1.78 (0.2)	2.02 (0.32)	2.1 (0.23)

Standard errors in parenthesis.

Table 3.14 - Ten year Oaxaca-Blinder twofold decomposition (overall population)

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
overall											
Natives	1.752***	1.791***	1.809***	1.852***	1.900***	1.958***	1.991***	2.024***	2.048***	2.069***	2.097***
	(0.00481)	(0.00406)	(0.00412)	(0.00404)	(0.00403)	(0.00432)	(0.00445)	(0.00444)	(0.00462)	(0.00468)	(0.00468)
Migrants	1.835***	1.848***	1.901***	1.869***	1.903***	1.944***	1.966***	1.993***	1.990***	1.989***	2.037***
	(0.0292)	(0.0205)	(0.0218)	(0.0231)	(0.0207)	(0.0195)	(0.0182)	(0.0165)	(0.0177)	(0.0160)	(0.0153)
difference	-0.0826**	-0.0578**	-0.0920***	-0.0171	-0.00315	0.0144	0.0242	0.0310	0.0585**	0.0795***	0.0593***
	(0.0296)	(0.0209)	(0.0221)	(0.0234)	(0.0211)	(0.0200)	(0.0187)	(0.0171)	(0.0183)	(0.0167)	(0.0160)
explained	-0.142***	-0.132***	-0.158***	-0.127***	-0.109***	-0.130***	-0.135***	-0.138***	-0.111***	-0.120***	-0.111***
	(0.0145)	(0.0110)	(0.0120)	(0.0112)	(0.00999)	(0.0104)	(0.00989)	(0.00962)	(0.00923)	(0.00938)	(0.00870)
unexplained	0.0598*	0.0741***	0.0659**	0.109***	0.106***	0.144***	0.159***	0.169***	0.169***	0.200***	0.170***
	(0.0291)	(0.0204)	(0.0217)	(0.0226)	(0.0208)	(0.0193)	(0.0188)	(0.0177)	(0.0185)	(0.0171)	(0.0163)
explained											
Education	-0.154***	-0.163***	-0.189***	-0.165***	-0.151***	-0.145***	-0.152***	-0.163***	-0.145***	-0.155***	-0.152***
	(0.0143)	(0.0111)	(0.0118)	(0.0110)	(0.00995)	(0.0101)	(0.00996)	(0.00989)	(0.00908)	(0.00926)	(0.00874)
Experience	-0.00139	0.00532*	0.00587*	0.00658**	0.00530*	-0.00557**	-0.00612*	0.0000117	0.00150	-0.00150	0.00193
	(0.00246)	(0.00218)	(0.00228)	(0.00247)	(0.00221)	(0.00202)	(0.00253)	(0.00255)	(0.00216)	(0.00251)	(0.00240)
Tenure	0.0261***	0.0282***	0.0310***	0.0358***	0.0420***	0.0332***	0.0350***	0.0377***	0.0406***	0.0446***	0.0439***
	(0.00391)	(0.00319)	(0.00310)	(0.00320)	(0.00295)	(0.00284)	(0.00301)	(0.00306)	(0.00322)	(0.00335)	(0.00308)
Married	-0.00103	0.000145	-0.00249*	0.000275	-0.00114	0.000762	0.000824	-0.00107	-0.00137	0.000654	-0.000366
	(0.00113)	(0.00107)	(0.00118)	(0.00130)	(0.000735)	(0.00128)	(0.00135)	(0.00105)	(0.000980)	(0.00116)	(0.00107)
Female	-0.0121**	-0.00217	-0.00306	-0.00433	-0.00482	-0.0129***	-0.0127***	-0.0112***	-0.00590*	-0.00888***	-0.00489**
	(0.00384)	(0.00328)	(0.00313)	(0.00273)	(0.00274)	(0.00306)	(0.00235)	(0.00257)	(0.00235)	(0.00231)	(0.00186)
unexplained											
Education	0.322*	0.227*	0.280**	0.151	0.160	0.0923	0.0759	0.178*	0.315***	0.258**	0.256**
	(0.143)	(0.0939)	(0.0979)	(0.123)	(0.0977)	(0.0880)	(0.0917)	(0.0881)	(0.0954)	(0.0800)	(0.0885)
Experience	0.0760	-0.00164	-0.00332	-0.0480	0.0250	0.00152	-0.0135	-0.00829	0.0619	-0.00502	0.00221
	(0.0747)	(0.0383)	(0.0444)	(0.0534)	(0.0392)	(0.0460)	(0.0433)	(0.0394)	(0.0394)	(0.0344)	(0.0331)
Tenure	-0.00822	-0.0232	-0.0173	-0.0306*	-0.0544**	-0.0560***	-0.0438**	-0.0598***	-0.0522***	-0.0424***	-0.0547***
	(0.0205)	(0.0139)	(0.0147)	(0.0141)	(0.0174)	(0.0123)	(0.0135)	(0.0113)	(0.0157)	(0.0120)	(0.0143)
Married	0.0311	0.0474	0.0477	-0.00249	0.0153	0.0283	0.0447*	0.0240	0.0150	0.0333	0.0164
	(0.0374)	(0.0247)	(0.0284)	(0.0265)	(0.0224)	(0.0233)	(0.0206)	(0.0211)	(0.0224)	(0.0189)	(0.0178)
Female	-0.0349	-0.0819***	-0.0418	-0.0228	-0.0802***	-0.0365	-0.0311	-0.0217	-0.0149	-0.0394*	-0.00511
	(0.0304)	(0.0232)	(0.0240)	(0.0235)	(0.0221)	(0.0188)	(0.0170)	(0.0164)	(0.0186)	(0.0160)	(0.0158)
Constant	-0.326	-0.0939	-0.200	0.0623	0.0409	0.115	0.127	0.0565	-0.156	-0.00513	-0.0444
	(0.190)	(0.114)	(0.119)	(0.158)	(0.118)	(0.118)	(0.119)	(0.113)	(0.119)	(0.0985)	(0.111)
N	10326	13949	14200	14189	13964	12946	13512	13457	12699	12496	12136

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.15 - Ten year Oaxaca-Blinder twofold decomposition (male population)

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
overall											
Natives	1.857***	1.901***	1.914***	1.940***	1.993***	2.068***	2.076***	2.125***	2.136***	2.154***	2.174***
	(0.00833)	(0.00670)	(0.00672)	(0.00674)	(0.00662)	(0.00706)	(0.00757)	(0.00748)	(0.00764)	(0.00775)	(0.00754)
Migrants	1.868***	1.845***	1.957***	1.944***	1.893***	1.989***	1.994***	2.056***	2.046***	2.014***	2.101***
	(0.0464)	(0.0354)	(0.0371)	(0.0324)	(0.0317)	(0.0284)	(0.0263)	(0.0255)	(0.0282)	(0.0233)	(0.0229)
difference	-0.0106	0.0556	-0.0430	-0.00444	0.100**	0.0792**	0.0815**	0.0692**	0.0902**	0.140***	0.0735**
	(0.0472)	(0.0360)	(0.0377)	(0.0331)	(0.0323)	(0.0293)	(0.0274)	(0.0266)	(0.0292)	(0.0245)	(0.0241)
explained	-0.104***	-0.102***	-0.156***	-0.133***	-0.0806***	-0.100***	-0.0983***	-0.128***	-0.103***	-0.0956***	-0.101***
	(0.0218)	(0.0168)	(0.0189)	(0.0175)	(0.0149)	(0.0143)	(0.0139)	(0.0145)	(0.0141)	(0.0140)	(0.0133)
unexplained	0.0931*	0.157***	0.113**	0.129***	0.181***	0.180***	0.180***	0.197***	0.194***	0.236***	0.175***
	(0.0445)	(0.0335)	(0.0361)	(0.0322)	(0.0314)	(0.0279)	(0.0272)	(0.0273)	(0.0294)	(0.0249)	(0.0246)
explained											
Education	-0.133***	-0.131***	-0.194***	-0.180***	-0.142***	-0.138***	-0.138***	-0.158***	-0.152***	-0.141***	-0.161***
	(0.0204)	(0.0164)	(0.0188)	(0.0175)	(0.0143)	(0.0141)	(0.0141)	(0.0148)	(0.0137)	(0.0134)	(0.0137)
Experience	-0.0105	-0.00418	0.000801	0.00580	0.00853	-0.00711	-0.00572	-0.00583	0.00161	-0.00444	0.00827*
	(0.00562)	(0.00444)	(0.00470)	(0.00521)	(0.00462)	(0.00388)	(0.00446)	(0.00429)	(0.00375)	(0.00390)	(0.00395)
Tenure	0.0428***	0.0367***	0.0445***	0.0437***	0.0554***	0.0393***	0.0455***	0.0368***	0.0519***	0.0458***	0.0498***
	(0.00680)	(0.00550)	(0.00528)	(0.00575)	(0.00502)	(0.00486)	(0.00549)	(0.00483)	(0.00553)	(0.00545)	(0.00488)
Married	-0.00307	-0.00310	-0.00678	-0.00301	-0.00226	0.00498	-0.000313	-0.000868	-0.00478	0.00420	0.00154
	(0.00433)	(0.00434)	(0.00356)	(0.00363)	(0.00263)	(0.00372)	(0.00304)	(0.00282)	(0.00318)	(0.00305)	(0.00211)
unexplained											
Education	-0.0113	0.0158	0.106	0.278	0.101	0.0329	-0.144	0.117	0.236	0.214	0.194
	(0.199)	(0.141)	(0.151)	(0.162)	(0.139)	(0.120)	(0.134)	(0.117)	(0.152)	(0.114)	(0.133)
Experience	0.179	0.0922	-0.00654	0.0258	0.112*	0.116	0.0102	0.0266	0.0896	0.0852	-0.0183
	(0.119)	(0.0674)	(0.0713)	(0.0756)	(0.0569)	(0.0668)	(0.0594)	(0.0599)	(0.0632)	(0.0520)	(0.0531)
Tenure	0.00608	-0.0422	-0.0230	-0.0352	-0.0613*	-0.0845***	-0.0525**	-0.0853***	-0.0566*	-0.0460*	-0.0552**
	(0.0324)	(0.0220)	(0.0182)	(0.0201)	(0.0259)	(0.0188)	(0.0188)	(0.0162)	(0.0274)	(0.0190)	(0.0208)
Married	0.102	0.138**	0.0812	0.0443	0.0452	0.0620	0.101**	0.100**	0.0447	0.0863**	0.0571*
	(0.0630)	(0.0448)	(0.0489)	(0.0467)	(0.0347)	(0.0340)	(0.0312)	(0.0355)	(0.0384)	(0.0271)	(0.0280)
Constant	-0.182	-0.0466	-0.0452	-0.184	-0.0162	0.0530	0.265	0.0381	-0.120	-0.104	-0.00322
	(0.277)	(0.181)	(0.177)	(0.193)	(0.156)	(0.161)	(0.170)	(0.146)	(0.188)	(0.137)	(0.165)
N	4117	5836	5849	5971	5875	5400	5770	5727	5478	5383	5245

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.16 - Ten year Oaxaca-Blinder twofold decomposition (female population)

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
overall											
Natives	1.684***	1.711***	1.735***	1.789***	1.833***	1.881***	1.929***	1.951***	1.983***	2.005***	2.038***
	(0.00563)	(0.00486)	(0.00502)	(0.00485)	(0.00490)	(0.00524)	(0.00524)	(0.00524)	(0.00558)	(0.00565)	(0.00583)
Migrants	1.805***	1.851***	1.858***	1.808***	1.912***	1.900***	1.937***	1.933***	1.940***	1.965***	1.981***
	(0.0369)	(0.0242)	(0.0259)	(0.0324)	(0.0276)	(0.0268)	(0.0251)	(0.0210)	(0.0219)	(0.0220)	(0.0204)
difference	-0.121**	-0.139***	-0.123***	-0.0188	-0.0789**	-0.0191	-0.00772	0.0185	0.0428	0.0402	0.0574**
	(0.0374)	(0.0247)	(0.0264)	(0.0327)	(0.0280)	(0.0273)	(0.0256)	(0.0217)	(0.0226)	(0.0227)	(0.0212)
explained	-0.151***	-0.158***	-0.159***	-0.115***	-0.121***	-0.127***	-0.143***	-0.128***	-0.109***	-0.125***	-0.107***
	(0.0192)	(0.0141)	(0.0146)	(0.0131)	(0.0130)	(0.0140)	(0.0135)	(0.0118)	(0.0115)	(0.0121)	(0.0108)
unexplained	0.0295	0.0188	0.0359	0.0960**	0.0417	0.108***	0.135***	0.146***	0.151***	0.166***	0.165***
	(0.0389)	(0.0256)	(0.0266)	(0.0318)	(0.0277)	(0.0270)	(0.0261)	(0.0231)	(0.0233)	(0.0236)	(0.0217)
explained											
Education	-0.166***	-0.185***	-0.184***	-0.149***	-0.154***	-0.149***	-0.163***	-0.163***	-0.138***	-0.166***	-0.142***
	(0.0198)	(0.0151)	(0.0151)	(0.0137)	(0.0137)	(0.0143)	(0.0138)	(0.0129)	(0.0120)	(0.0125)	(0.0110)
Experience	0.0000967	0.00571*	0.00417	0.00346	0.00126	-0.00737**	-0.00858**	-0.00220	-0.00216	-0.00325	-0.00503
	(0.00204)	(0.00230)	(0.00250)	(0.00234)	(0.00226)	(0.00250)	(0.00300)	(0.00331)	(0.00271)	(0.00351)	(0.00311)
Tenure	0.0153***	0.0215***	0.0220***	0.0297***	0.0319***	0.0291***	0.0274***	0.0382***	0.0320***	0.0439***	0.0405***
	(0.00465)	(0.00372)	(0.00368)	(0.00361)	(0.00350)	(0.00359)	(0.00342)	(0.00398)	(0.00382)	(0.00430)	(0.00401)
Married	0.0000347	-0.000162	-0.000559	0.000676	0.000114	-0.000129	0.00139	-0.000529	-0.0000435	-0.000295	-0.00103
	(0.000219)	(0.000288)	(0.000544)	(0.000695)	(0.000351)	(0.000631)	(0.00115)	(0.000613)	(0.000148)	(0.000680)	(0.00110)
unexplained											
Education	0.669***	0.449***	0.475***	-0.0251	0.203	0.147	0.330**	0.243	0.406***	0.287*	0.331**
	(0.198)	(0.122)	(0.119)	(0.184)	(0.142)	(0.128)	(0.115)	(0.136)	(0.119)	(0.113)	(0.117)
Experience	0.0417	-0.0806	0.0119	-0.119	-0.0517	-0.115	-0.00658	-0.0456	0.0406	-0.0806	0.0297
	(0.0915)	(0.0453)	(0.0551)	(0.0757)	(0.0545)	(0.0648)	(0.0611)	(0.0549)	(0.0489)	(0.0459)	(0.0415)
Tenure	-0.0162	0.00364	-0.0148	-0.0228	-0.0477*	-0.0322*	-0.0340	-0.0334*	-0.0493*	-0.0378*	-0.0537**
	(0.0270)	(0.0158)	(0.0213)	(0.0204)	(0.0225)	(0.0164)	(0.0198)	(0.0152)	(0.0192)	(0.0156)	(0.0197)
Married	-0.00992	-0.00600	0.0277	-0.0317	-0.00635	0.0130	0.00501	-0.0297	0.00125	-0.0141	-0.0157
	(0.0431)	(0.0281)	(0.0349)	(0.0317)	(0.0303)	(0.0319)	(0.0278)	(0.0254)	(0.0267)	(0.0272)	(0.0229)
Constant	-0.655*	-0.347*	-0.464**	0.295	-0.0557	0.0952	-0.159	0.0124	-0.247	0.0109	-0.127
	(0.259)	(0.144)	(0.149)	(0.256)	(0.176)	(0.174)	(0.165)	(0.172)	(0.147)	(0.148)	(0.146)
N	6209	8113	8351	8218	8089	7546	7742	7730	7221	7113	6891

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Bibliography

Anees, M, Sajjad, M and Ahmed, I, (2011), *A Counterfactual Decomposition Analysis of Immigrants-natives Earnings in Malaysia*, Economics Discussion Papers, No 2011(51), Kiel Institute for the World Economy - http://www.economicsejournal.org/economics/discussionpapers/2011-51

Arrow, K,J, (1998), *What has Economics to Say about Racial Discrimination*, Journal of Economic Perspectives, Vol. 71(5), 91-100

Barrera-Osorio, Felipe, Vincente Garcia-Moreno, Harry Anthoney Patrinos, and Emilio Porta. (2011). *Using the Oaxaca-Blinder Decomposition Technique to Analyze Learning Outcomes Changes over Time: An Application to Indonesia's Results in PISA Mathematics. World Bank* Policy Research Working Paper 5584

Barret, A & O'Connell, P.J (2000), *Is There a Wage Premium for Returning Irish Migrants,* The Economic and Social Review, Economic and Social Studies, Vol. 32(1), 1-21

Barret, A & O'Connell, P.J (2000), *Is There a Wage Premium for Returning Irish Migrants,* The Economic and Social Review, Economic and Social Studies, Vol. 32(1), 1-21

Becker, G, (1957), *The Economics of Discrimination*, The University of Chicago Press,

Becker, G, (1962), *Investment in Human Capital: A Theoretical Analysis*, Journal of Political Economy, The University of Chicago Press, Vol.70(5), 9-49

Biltagy, M (2014), *Estimation of Gender Wage Differentials in Egypt using Oaxaca Decomposition Technique,* Topics in Middle Eastern and African Economies Vol. 16(1), 17-42

Becker, G, (1973), *The Economics of Discrimination*, The University of Chicago Press, Chicago

Bevelander, P & Nielson, H, S (2001), *Declining Employment Success of Immigrant Males in Sweden: Observed or Unobserved Characteristics?*, Journal of Population Economics, Vol.14(3), 455-471

Blinder, A (1973), *Wage Discrimination: Reduced Form and Structural Estimates*, The Journal of Human Resources, Vol.4(8),436-455

Severin, B (2012) "What Happened to Airline Market Power?" *University of California Berkeley Haas School of Business working paper* (2011) available at http://faculty.haas.berkeley.edu/borenste/AirMktPower2013.pdf

Borjas,G (1987), *Self-Selection and the Earnings of Immigrants*, American Economic Review, Vol. 77, 531-553

Borjas, G (1989), *Economic Theory and International Migration*, International Migration Review, Vol.23(3)

Borjas, G (1995), *The Economic Benefits from Immigration*, The Journal of Economic Perspectives, Vol. 9(2), 3-22

Borjas, G & Bratsberg, B (1996), *Who Leaves? The Outmigration of the Foreign Born,* The Review of Economics and Statistics, Vol. 78(1), 165-176

Brown, R, S, Moon, M, Zoloth, B ,S (1980*), Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials*, Journal of Human Resources, Vol.15(1), 3-28

Charles, K, K & Guryan, J, (2008), *Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination,* Journal of Political Economy, Vol.116(5), 773-809

Charles, K, K, & Guryan, J, (2008), *Prejudice and Wages: An Empirical Assessment of Becker's "The Economics if Discrimination",* Journal of Political Economy, Vol.115(5), 773-806

Chiswick, B.R (1999), *Are Immigrants favourably Self-Selected?*, The American Economic Review, Vol. 89(2), 181-185

Co, Yun & Gang (1998), *Returns to Returning: Who Went Abroad and What Does it Matter*, IZA Discussion Paper No.19

Cooke, T, J (2011), *It is not Just the Economy: Declining Migration and the Rise of Secular Rootedness*, Population, Space and Place, Vol.17(3) 193-203

Constant, A & Massey, D.S (2002), *Self-Selection, Earnings and Out-Migration: A Longitudinal Study of Immigrants to Germany*. IZA Discussion Paper No.672

Cotton, J (1988), *On the Decomposition of Wage Differentials*, Review of Economics and Statistics,Vol.10(12), 236-243

de Coulon, A & Piracha, M (2004), *Self-Selection and the Performance of Return Migrants: A Source Country Perspective*, Journal of Population Economics, Springer, Vol. 18(4), pages 779-807

DiNardo, J, et al (1996), *Labour Market Institutions and the Distribution of Wages, 1973-1992: A Semi-Parametric Approach*, Econometrica, Vol. 64(5), 1001-1044

Dustmann, C (2003), *Return Migration, Wage Differentials, and the Optimal Migration Duration*, European Economic Review, Vol. 45(2), 353-369

Dustmann, C, Fabbri, F & Preston, I (2005), *The Impact of Immigration on the British Labour Market*, The Economic Journal, Vol.115(507), F324-F341

Dustmann, C & Weiss, Y (2007), *Return Migration: Theory and Empirical Evidence from the UK*, British Journal of Industrial Relations, Vol. 45(2), 236-256

Entorf,. H & Tatsi, E (2009), Migrants at School: Educational Inequality and Social Interaction in the UK and Germany, *IZA Discussion Paper No. 4175*

The Failed States Index (2011), 02/05/17, http://library.fundforpeace.org/fsi11

Fortin, N ,M, & Lemieux, T (1998), *Rank Regressions, Wage Distributions, and the Gender Gap*, Journal of Human Resources, Vol.33(3), 610-643

Heckman, J.J & Honore, Bo.E (1990), *The Empirical Content of the Roy Model*, Econometrica, Vol.58(5), 1121-1149

Jann, B (2008), *A Stata Implementation of Blinder-Oaxaca Decomposition*, The Stata Journal, Vol. 8(4), 453-479

Jasso, G & Rosenzweig, M.R (1990), *Self-Selection and the Earnings of Immigrants: Comment*, American Economic Review, Vol. 80(1), 298-304

Jimenez-Rubio, J & Hernandez-Quevedo, C, (2010), *Inequalities in the use of health services between immigrants and the native population in Spain: what is driving the differences?*, European Journal of Health Economics, Vol. 12(1), 17-28

Lee, L.F (1978), *Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependant Variables*, International Economic Review, Vol.19(2), 415-433

Macado, J & Mata, J (2005), *Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression*, Journal of Applied Econometrics, Vol. 20(4), 445-465

McElroy, M.B & Horney, M.J (1981), *Nash-Bargained Household Decisions; Toward a Generalization of the Theory of Demand*, International Economic Review, Vol. 22(2), 333-349

Mincer, J & Polachek, S (1974), *Family Investment in Human Capital: Earnings of Women*, Journal of Political Economy, Vol.82(2), 76-110

Miller, R.A (1984), *Job Matching and Occupational Choice*, Journal of Political Economy, Vol. 92(6), 1086-1120

Miller, P.W (1987), *The Wage Effect of the Occupational Segregation of Women in Britain*, The Economic Journal, Vol. 97(388), 885-896

Miller, C (1993), *Actual Experience, Potential Experience or Age, and Labour Force Participation by Married Women*, American Economics Journal, Vol.21(4), 60-66

Neumark,D (1988), *Employers' Discriminatory Behaviour and the Estimation of Wage Discrimination*, The Journal of Human Resources, Vol.23(3), 279-295

Nichols, A & Schaffer, M (2007), Clustered Errors in Stata, file://homefs03/home/vfj1/cluster%20robust.pdf

Nickell, S & Saleheen, J (2009), *The Impact of Immigration on Occupational Wages: Evidence from Britain,* Spatial Economics Research Centre, SERC Discussion Paper 2009

Oaxaca, R (1973), *Male-Female Wage Discrimination in Urban Labour Markets*, International Economic Review, Vol.14(3), 693-709

Oaxaca, R & Ransom, M, R (1994), On discrimination and the decomposition of wage differentials, Journal of Econometrics, Vol.61(1), 5-21

Office for National Statistics. Social Survey Division. (2015). *Quarterly Labour Force Survey Household Dataset, January - March, 2015*. [data collection]. UK Data Service. SN: 7816, http://dx.doi.org/10.5255/UKDA-SN-7816-1**.**

Phipps, S, Burton, P, Lethbridge, L, *In and Out of the Labour Market: Long-term income consequences of Child-related Interruptions to Women's Paid Work,* Canadian Journal of Economics, Vol.34(2), 411-429

Polavieja, J, G (2004), *Task Specificity and the Gender Wage Gap*, Sociology Working Papers, University of Oxford, Paper Number 2004-03

Quheng, D (2007), *Earnings Differential between Urban Residents and Rural Migrants: Evidence from Oaxaca-Blinder and Quantile Regression Decompositions,* Chinese Journal of Population Science, 2007-02

Ramos, F.A, (1992), *Out-Migration and Return Migration of Puerto Ricans,* NBER Working Paper 6905

Reimers, C, W (1983), *Labour Market Discrimination Against Hispanic and Black Men*, The Review of Economics and Statistics, Vol.65(4), 570-579

Robinson, C & Tomes, N, (1982), *Self-Selection and Interprovincial Migration in Canada*, The Canadian Journal of Economics, Vol.15(3), 474-502

Rooth, A.D, & Saarela, J, (2006), *Selection in Migration and Return Migration: Evidence from Micro Data,* Economic Letters, Vol. 94(1), 90-95

Roy, A.D, (1951), *Some Thoughts on the Distributions of Earnings*, Oxford Economic Papers – New Series, Vol. 3(2), 135-146

Ruhs, M & Vargas-Silva, C (2015), *Briefing: The Labour Market Effects of Immigration*, available at http://www.migrationobservatory.ox.ac.uk/sites/files/migobs/Briefing%20-%20Labour%20Market%20Effects%20of%20Immigration_0.pdf

Sjaastad, L.A, (1962), *Costs and Returns of Human Migration*, Journal of Political Economy, Vol.23(3), 457-485

StataCorp (2013), *Stata Statistical Software: Release 13*, College Station, TX: StataCorp LP

StataCorp (2013), *Labour Force Survey: User Guide – Volume 3 – Details of LFS Variables 2013 – Version 4 February 2014*, College Station, TX: StataCorp LP

Suh, J (2009), *Decomposition of the change in the gender wage gap*, Research in Business and Economics Journal, Research in Business and Economics Journal Vol.1, 1-18

Woodcock, S (2008), *Wage Differentials in the Presence of Unobserved Worker, Firm and Match Heterogeneity*, Labour Economics, Vol.15(4), 771-793

Willis, R.J, & Rosen, S (1979), *Education and Self-Selection*, Journal of Political Economy, Vol.87(5), S7-36

Wooldridge, J, M (2003), *Cluster-Sample Methods in Applied Econometrics,* The American Economic Review, Vol. 93(2), 133-138

Yun, M.S (2007), *A Simple Solution To The Identification Problem In Detailed Wage Decompositions*, Economic Inquiry, Vol.43(4), 766-772

Zaiceva, A & Zimmerman, K , *Returning Home in times of Trouble? Return Migration of EU Enlargement Migrants During the Crisis*, IZA Working Paper 7111